Designing cooperative interaction of automated vehicles with other road users in mixed traffic environments

interACT D.2.1 Preliminary description of psychological models on human-human interaction in traffic

Work package	WP2: Psychological Models on Human Interaction and Intention Recognition Algorithms
Task	Task 2.1: Naturalistic, cross-cultural observation of present human-human interactions
Authors	**Dietrich, André (TUM)**, Bengler, Klaus (TUM), Evangelia Portouli (ICCS), Nathanael, Dimitris (ICCS), Ruenz, Johannes (BOSCH), Wu, Jingyuan (BOSCH), Merat, Natasha (ITS LEEDS), Madigan, Ruth (ITS LEEDS), Lee, Yee Mun (ITS LEEDS), Markkula, Gustav (ITS LEEDS), Giles, Oscar (ITS LEEDS), Fox, Charles (ITS LEEDS), Camara, Fanta (ITS LEEDS)
Dissemination level	Public (PU)
Status	Final
Due date	31/05/2018
Document date	31/05/2018
Version number	1.0
	This work is part of the interACT project. interACT has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395. Content reflects only the authors' view. The Innovation and Networks Executive Agency (INEA) is not responsible for any use that may be made of the information it contains.

This is a draft version of deliverable D2.1
which has not been approved by the EC, yet.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Table of contents

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Glossary of terms

Term	Description
Vulnerable Road User (VRU)	Road users with a higher fatality rate per accident than other groups. In particular, pedestrians, bicycles, motorised two-wheelers and non-motorised traffic.
Mixed Traffic	Usually referred to traffic consisting of different types of road users (such as pedestrians, busses, cars, etc.). Also used in the context of traffic consisting of automated vehicles mixed and human road users.
Transition Phase	Projected or theoretical time frame between the first vehicles with higher automated driving functions (SAE3+) being integrated into traffic and the majority of motorized traffic being automated.
Reaction	Reaction [of one road user to other road users]: Road user A is said to have reacted to road user B if A's behaviour can be interpreted as A having perceived B and A's behaviour having been affected to some extent by B.
Interaction	Interaction [between road users]: Road users A and B are said to be interacting if they are both reacting to one another (by the above definition of reaction).
Traffic Conflict	An observable situation in which two or more road users approach each other in space and time to such an extent that a collision is imminent if their movements remain unchanged.
Edging	Moving forward with very low velocity usually to indicate a desired trajectory (e.g. turning). Edging is mostly used by drivers, trying to pull out of a parking space with limited vision or while turning on congested priority lanes.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

List of abbreviations and acronyms

Abbreviation	Meaning
AV	Automated Vehicle
RU	Road User
HRU	Human Road User
VRU	Vulnerable Road User
HMI	Human Machine Interface
D	Deliverable
WP	Work package
eHMI	External Human-Machine-Interface of the AV that is meant to communicate with surrounding traffic participants
DDT	Dynamic Driving Task
SAE	SAE International (initially established as Society of Automotive Engineers)
OEDR	Object and Event Detection and Response
MA	Movement Achieving
MS	Movement Signalling
PA	Perception Achieving
PS	Perception Signalling
RUBQ	Adolescent Road User Behaviour Questionnaire
MOT	Multiple Object Tracking
MAP	Maximum A Posteriori
HOG	Histogram of Oriented
SVM	Support Vector Machines
TTC	Time To Collision
MCMC	Markov Chain Monte Carlo
SMC	Sequential Monte Carlo
ABC	Approximate Bayesian Computation
CCPU	Coordination and Communication Planning Unit: interACT central software unit that plans AV behaviour and explicit HMI control in an integrated, timely, and synchronised manner

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Executive Summary

Automated Vehicles (AVs) have seen rapid technological development over the last decade and will soon be deployed on public roads. However, road traffic is unlikely to become fully automated in the near future. Instead, AVs will share the road space with human road users (HRUs), including cyclists, pedestrians and drivers. A major challenge in the development of AVs is understanding how these vehicles should interact with HRUs to ensure safe and efficient traffic flow. InterACT aims to understand how interactions unfold between road users, in order to ensure the safe integration of AVs into mixed traffic environments.

To safely navigate mixed traffic environments (traffic environments used by both AVs and HRUs), AVs will need to behave in a way, which HRUs can readily anticipate. If AVs show unexpected non-human like behaviour, this may lead to traffic flow inefficiencies or even increase the risk of traffic accidents. In addition, AVs will need to anticipate and understand the actions of HRUs when planning how to move through the environment. Thus, studying how HRUs interact with each other in complex traffic environments is of critical importance for the success of AVs.

This deliverable provides an overview of Work Package 2 (WP2) from the interACT project. WP2 has two primary goals. Firstly, it aims to define precise terminology for describing interactions between HRUs. Secondly, it provides an overview of a large cross-cultural observation study examining how HRUs interact in complex traffic environments. The study was carried out in three major European cities: Leeds (UK), Athens (Greece) and Munich (Germany). Thus, the study provides insights into cross-cultural differences in HRUs behaviours. The main results show that HRUs try to avoid communicating explicitly with each other (e.g. by using hand gestures or flashing headlights) and use the kinematic information of other road users to plan their next actions. Only in congested traffic situations, additional means to communicate intent were observed.

This document serves as an input for WP3 “Cooperation and Communication Planning Unit” by detailing which situations require interaction and for WP4 “Suitable HMI for successful human-vehicle interaction” by depicting how interactive situations between different road users progress over time and what kind of communication is used by HRUs.

WP2 and the data analysis from the observations are still ongoing, more elaborate models of interaction will be derived from the data and reported within the deliverable D2.2.

This report is part of interACT project that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 723395.

1. Introduction

1.1 Purpose and scope

Urban traffic is complex. Unique road layouts coupled with a variety of different traffic participants and their individual interpretation of traffic regulations create an endless number of possible situations that human drivers have to cope with nowadays. For a safe integration into urban traffic and a minimal amount of interventions required by the driver, automated vehicles ideally will have to deal with these situations at least as good as a human driver.

Following the SAE J3016 (2016), the Dynamic Driving Task (DDT) is categorized into the following categories:

- Dynamic Motion control (longitudinal and lateral),
- Object and Event Detection and Response (OEDR),
- Manoeuvre planning and
- Enhancing conspicuity (gestures, turn indicator etc.).

If a vehicle is supposed to be driven in higher modes of automation, all of these categories have to be performed by the AV. While controlling the vehicle's motion control in regards to the navigational destination is straightforward to understand, as street regulations can be used as a foundation, OEDR and conspicuity are highly subjective, especially when encountering other road users. Therefore, for a safe integration of AVs into urban traffic it is essential to understand "normal" and "interaction requiring situations" in different use cases in current urban traffic

Naturalistic observation studies were conducted within WP2 to understand human-human interactions in current urban traffic. These observations were simultaneously carried out in Athens (Greece), Leeds (UK) and Munich (Germany) to observe cross-cultural effects.

While each traffic situation in urban traffic is somewhat unique and at times ambiguous, human road users (most of the time) manage to understand how to safely reach a destination – even if the location of the traffic situation is unknown. This means that, human road users have some sort of mutual understanding and anticipation towards traffic that lets them resolve any potential **traffic conflict** (see Ch. 2.2.1).

The main purpose of this deliverable is to understand and model current traffic to help identifying **interaction-demanding situations** (see Ch. 2.2.1) and how traffic participants resolve them nowadays, using their available **means of communication** (see Ch. 6.1).

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

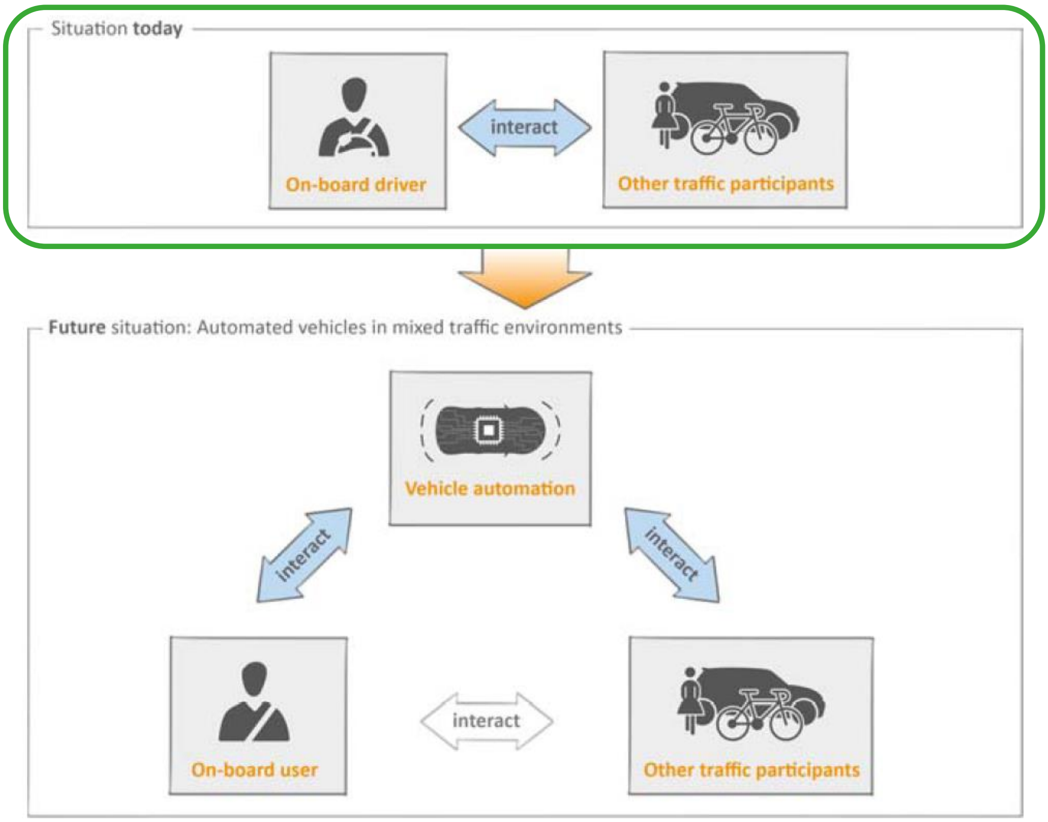


Figure 1: Scope of this deliverable within interACT: Understanding the traffic interaction today will enable implementing AVs onto urban environments, that are able to correctly assess any given situation, manoeuvring and communicating in an expectation conforming way around other traffic participants.

1.2 Intended Readership

This deliverable gives an insight into current urban traffic interactions in use cases defined in D1.2 of WP1. The results create a knowledge base for the development of automated vehicles, by observing and understanding situations, which the automated vehicle has to resolve by employing manoeuvres and explicit communication strategies to ensure safe encounters with other road users. Therefore, this document serves primarily as an input for **WP3** and **WP4** while also providing useful insights to **traffic researchers** by elaborating on different observation methods. The document is expected to support the discussion on cross-cultural differences regarding interaction in urban traffic between different research teams, including the twinning team of the NHTSA project AVIntent.

As this deliverable is public, it will also serve **everyone interested** in this topic. It depicts the complexities of modelling urban traffic scenarios and the challenges in general for introducing automated vehicles onto city roads.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

1.3 Relationship with other interACT deliverables

interACT System Architecture", as the observed locations in **Task 2.1** "Naturalistic, cross-cultural observation of present human-human interactions" are based on the theoretically defined use cases. Using communication concepts developed in **WP4** "Suitable HMI for successful human-vehicle interaction", their influence on the behaviour of other traffic participants is measured and modelled in **Task 2.2** "Development of human-human and human automation interaction models (qualitative and quantitative)".

The other key objective of WP2 is the development of real-time algorithms for improved sensor based intention recognition and path prediction for traffic participants that will be provided to **WP3** for the situation assessment module of the CCPU and the development of expectation-conforming algorithms and **WP5** to set up the demonstrators.

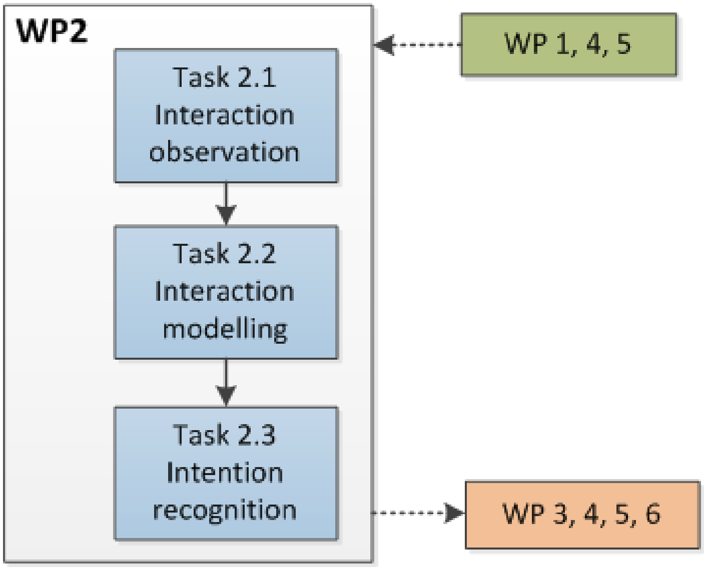


Figure 2: Orientation of Work Package 2 within interACT.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

2. Interaction in Traffic

2.1 Definition of Interaction

The concept of **interaction** is central to the interACT project, and it is therefore important to define what is, for the purposes of this project, meant by this term. Standard definitions include:

> Interaction*:* An occasion when two or more people or things communicate with or react to each other. (Cambridge Dictionary)

Where, further, the concept of a **reaction** is defined as:

> Reaction*:* Behaviour, a feeling or an action that is a direct result of something else. (Cambridge Dictionary)

Based on the definitions above, the following definitions are introduced here for the interACT project:

> **Reaction** [of one road user to other road users]: Road user A is said to have reacted to road user B if A's behaviour can be interpreted as A having perceived B and A's behaviour having been affected to some extent by B.

Note that by this definition, an act of communication from one road user to another (such as mentioned in the dictionary definition of interaction above) is a special case of a reaction, since the communication would not have occurred if the other road user were not present. Furthermore, note that the definition is neutral on whether or not A's perception and decision-making is "conscious" or not; i.e., A might be interpreted as reacting to B even if A is afterwards completely incapable of remembering B.

Furthermore:

> **Interaction** [between road users]: Road users A and B are said to be interacting if they are **both** reacting to one another (by the above definition of **reaction**).

It should be noted that by this definition, if just one road user reacts to another, this is not considered an interaction. Consider, for example, a situation where (1) a car driver passes a pedestrian on a sidewalk without changing speed and without giving noticeable visual attention to the pedestrian, and (2) the pedestrian waits for the car to pass before crossing the road. Thus, the pedestrian reacts to the car driver, but the car driver does not react to the pedestrian, hence this was not an interaction by the definition proposed above. However, please note that the definitions are by necessity not completely exact, and do leave room for interpretation. The example just given would be classified as an interaction if the car driver is nevertheless judged to have perceived the pedestrian at some level, and to subsequently not look further at the pedestrian as part of a strategy (conscious or not) to make it clear to the pedestrian that the car driver will not be yielding.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

2.2 Towards a theory of interactions in traffic

This section provides early sketches towards a conceptual theory to provide structure in thinking and reasoning about road traffic interactions, both between humans and between humans and AVs.

2.2.1 Traffic conflicts and interaction-demanding situations

Up until recently, much of the research into traffic interactions has focused on situations where a collision or near miss situation between two or more traffic participants might arise, and there is almost 30 years of research investigating these so-called “traffic conflicts”. Hydén (1996) suggested the following definition:

> **Traffic conflict**: An observable situation in which two or more road users approach each other in space and time to such an extent that a collision is imminent if their movements remain unchanged.

Actual near miss situations and traffic collisions are relatively rare events in real traffic, but it is nevertheless interesting to note that all of the interACT use cases defined in Deliverable 1.1 essentially relate to the concept of collision avoidance. In other words, also completely non-critical, routine interactions in traffic arise because two or more actors are competing for the same location in space, and need to determine order of access to this location.

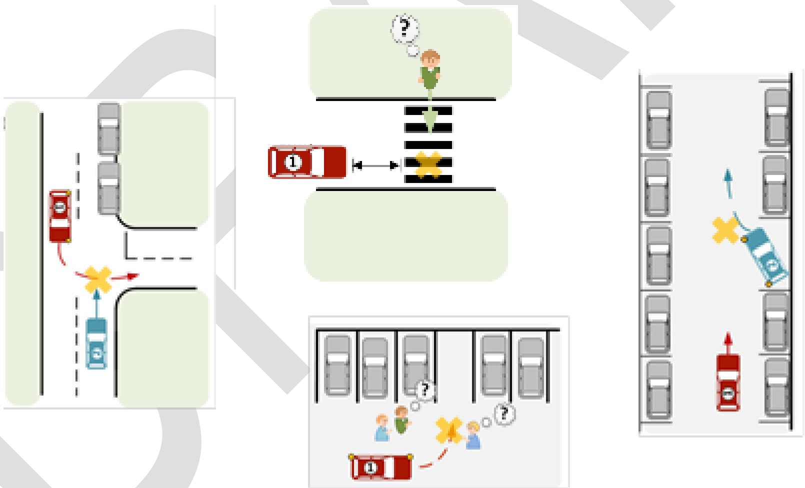


Figure 3: Illustration of how use cases considered in interACT can fundamentally be understood as two or more road users trying to determine the order of access to a region of space (to avoid collision at the location in question).

This report is part of interACT project that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 723395.

As a generalisation of the concept of a "traffic conflict" as defined above, it might therefore be useful to introduce the following concept:

> **Interaction-demanding situation**[1] [in road traffic]**:** An observable situation in which two or more road users are positioned, and/or moving, and/or explicitly communicating, relative to each other in space and time in such a way that the road users and/or a third party observer are likely to interpret the road users to be **intending or wanting to occupy the same region of space at the same time.**

Note that by the definition above, most or all "traffic conflicts", i.e., situations with an established objective collision course, will also be "interaction-demanding situations", but this new concept extends beyond these situations, to connect also with phenomena such as road user intentions/goals, communication, human perception, and human interpretation of traffic situations. This seems reasonable, under the assumption that all of these phenomena will be part of determining human behaviour in traffic interactions. An interaction-demanding situation may arise even in the absence of an objective collision course, for example if road user A is uncertain about road user B's intended movement path, or about whether road user B has correctly understood A's intentions, etc. Furthermore, note that an interaction-demanding situation is resolved when one or more of the road users have changed their behaviour in such a way that it is no longer likely that anyone will interpret the road users as intending to occupy the same region of space at the same time (analogously to how a traffic conflict is resolved when one or more of the road users have changed their behaviour such that there is no longer an objective collision course). For example, this may occur when one of the involved road users reacts to the interaction-demanding situation by changing their movement, and/or exhibits a communicative gesture, to indicate clear intent of yielding to the other road user(s).

The focus of interACT WP2 can be formulated as investigating the contexts in which interaction-demanding situations arise and how they get resolved, resulting in useful information about where AV interaction solutions are likely to be needed, and how these solutions should be implemented.

2.2.2 Types of interaction-relevant road user behaviour

The definition of "interaction" proposed further above depends heavily on the definition of "reaction", which in turn depends heavily on the concept of "road user behaviour". Based on existing knowledge of how people behave in traffic interactions (Risser, 1985; Sucha et al, 2017), combined with discussions so far in interACT about AV-human interactions, the following four main types of interaction-relevant behaviour are defined:

[1] To align better with Deliverable 1.1, one could use the term "interaction-demanding **scene**", but "situation" seems more appropriate as soon as one moves away from the specific and relatively technical context of defining technology use cases

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Movement-achieving (MA) behaviour: Behaviour that moves a road user in the world.

This definition applies to any human body or vehicle movement that has an effect of how the region of space occupied by a road user changes, or does not change, over time. This behaviour can typically be succinctly described in terms of positions, speeds, accelerations, etc.

Movement-signalling (MS) behaviour: Behaviour that can be interpreted as giving information on how a road user intends to move in the future.

An alternative term could be "intention-signalling behaviour". Examples include (1) a pedestrian walking in a way that can be taken to suggest that their current path is unlikely to change, or (2) a human-driven car or AV decelerating to yield to another road user, or (3) the same vehicle also showing an external sign indicating the intention to yield (e.g. headlights or some AV eHMI).

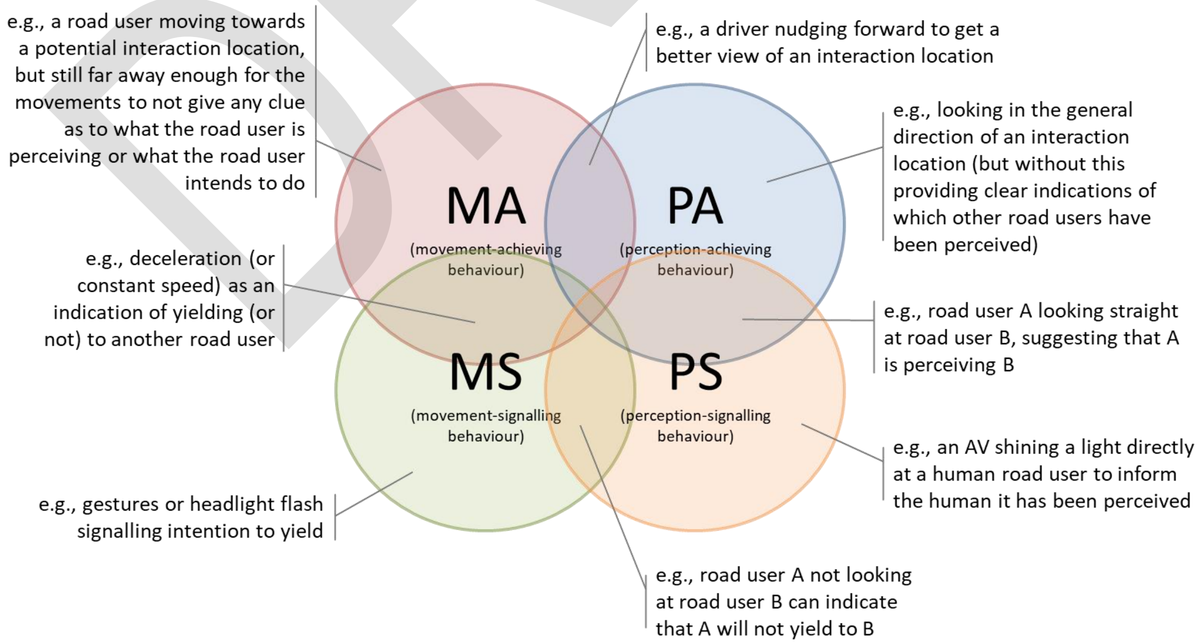
Perception-achieving (PA) behaviour: Behaviour that determines what a road user perceives.

This definition applies to any human body or vehicle movement that has an effect on what the road user perceives. Examples include head/eye movements, or a vehicle advancing in an intersection to get a better view of surrounding traffic.

Perception-signalling (PS) behaviour: Behaviour that can be interpreted as giving information on what a road user is perceiving.

Examples include (1) driver eye or head orientation/movement indicating that the driver is looking at a pedestrian while approaching a crossing, (2) a pedestrian head/arm posture indicating that the pedestrian is busy interacting with a mobile phone, (3) an AV shining a directed light at a certain human road user (in an attempt) to indicate that the AV has detected the human road user.

As should be clear from the above, these four types of behaviours are not mutually exclusive, rather the opposite; the figure below provides some examples of possible ways they may overlap.



This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Figure 4: An illustration of how four types of interaction-relevant road user behaviours relate to each other.

This terminology, with the four behaviour types, could be useful when describing traffic interactions and when considering how an AV should interact with human road users. It may be noted that, as illustrated in Figure 2, two other concepts that have been used previously in interACT can be rather neatly defined using these terms:

> **Implicit communication:** A behaviour which is at the same time **both achieving and signalling** movement and/or perception.

For example, decelerating signalling intention to yield, or looking at another road user signalling perception of that road user.

> **Explicit communication:** A behaviour signalling perception and/or movement without at the same time achieving either of these.

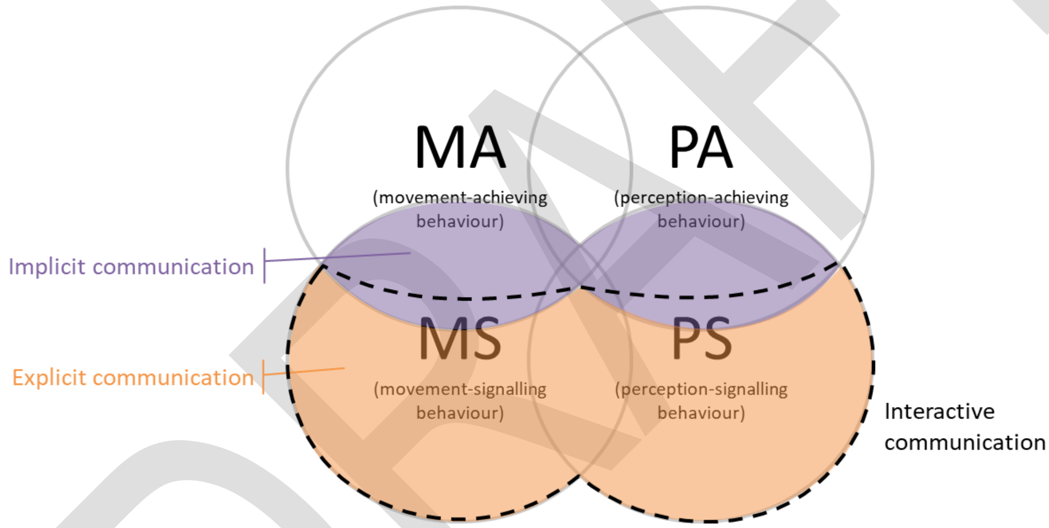


Figure 5: Illustration of the concepts of explicit and implicit communications, in terms of the four types of interaction-relevant road user behaviours defined here.

In other words, any gestures, signs, utterances, or other communicative acts that signal how one intends to move in the world without making use of one's actual movement in the world, or which signal perception without actually altering perception. In practice, any explicit communication of what one perceives or one's intended movement will typically be interpretable as being carried out with a specific intent to communicate, for example to initiate cooperation in an interaction. However, note that according to these definitions of implicit/explicit communication, an action like braking harder than normal to emphasise that one is yielding, is still considered implicit communication, even though it could be argued that it is performed with "intent". In the AV case this will in practice mean that implicit communication will always refer to how the AV's movements in the world gives indications as

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

to what it is doing. However, to accommodate the nuance relating to intentions towards cooperation and similar, we also propose the following definition:

> **Interactive communication:** A movement-signalling or perception-signalling behaviour that can be interpreted as being carried out with the specific goal of resolving an interaction-demanding situation.

Note that interactive communication, by this definition, covers communicative behaviour that aims for polite cooperation, such as slowing down or showing explicit indications of yielding, but just as well communicative behaviour that aims to assert one's own right of passage above that of others, such as honking or increasing speed.

2.3 Possible factors influencing interaction

The complexity of urban traffic does not only stem from the variety of road users but rather has a many different factors changing the outcome of a given situation.

Examples

A vehicle's acceleration from standing still in a scenario with no visual obstruction on a non-priority road on an intersection can likely be interpreted as a **Movement-achieving** behaviour as the vehicle is driving towards the intersection. If the drivers view is obstructed by parked vehicles, buildings or other influences, accelerating from a near standstill could indicate a **Perception-achieving** behaviour. While in the first case the vehicle will likely proceed its acceleration, in the second case it will likely do so after reassessing the situation from a new position.

Another example is jaywalking: in low dense traffic, pedestrians will usually pass after a vehicle if the gap to a following up vehicle is sufficiently large. Therefore, pedestrians expect that drivers will not change their **Movement-achieving** behaviour. However, if the traffic is congested on the very same road section, drivers might break early to let a pedestrian pass, as the own yielding does not impend the goal reaching. In this situation a driver perceives the pedestrian's **Perception and/or Movement-signalling** behaviour as well as the congestion in front, thus changing from **Movement-achieving** to **Movement-signalling** (i.e. decelerating) and **Perception-signalling** (e.g. flashing lights) behaviour.

While not critical, both examples show that outer environmental influences may affect the behaviour of one traffic participant, which can be expected by other traffic participants, thus heavily influencing a traffic encounter.

While some factors can be identified using available information (e.g. the road layout, traffic density, allowed velocities etc.), other, subtler influences might have big effects on the outcome of a pedestrian vehicle encounter.

A list of influencing factors was created, with over 40 potential influences identified. The potential impact of each factor was rated by each research team. Examples, which were rated as having the

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

highest impact on the observation, are traffic density, number of lanes and allowed traffic velocity as well as time of day, age and gender. As not all factors can be controlled in a naturalistic traffic observation, they were protocolled as possible independent variables.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

3. Observational studies in interACT[2]

3.1 Research questions & expected outcome

Leading up to the naturalistic observation study, research questions were formulated to create a common understanding of the individual aims of each partner. The overall research goal is *to identify invariant or persistent features of interaction between road users (RU), In order to build - computational / algorithmic RU interaction models to be implemented in (semi or fully) automatic vehicles*. Thus, the general aim can be transformed into a core question:

"What can we learn from today's interactions in traffic to improve future communication and interaction with automated vehicles?"

To address this question within WP2, research questions were formulated individually and grouped into overall topics giving an insight on the expected outcome of each partner while implying requirements for the methodology.

Table 1: Clustered Research Questions of each partner.

Overall Topic	Research Questions
General aim	[…] to identify invariant or persistent features of interaction between road users (RU), In order to build - computational / algorithmic RU interaction models to be implemented in (semi or fully) automatic vehicles.
	"What can we learn from today's interactions in traffic to improve future communication and interaction with automated vehicles"
"How do traffic participants interact?"	From what cues do drivers interpret a pedestrian's intention to cross? (and/or the other way around "what tacit or explicit signs do pedestrians use to communicate their intent?)
	From what cues do pedestrians interpret a driver's intent to give him priority or not? (and/or the other way around "what tacit or explicit signs do vehicle drivers use to communicate their intent to pedestrians or other drivers?)
	From what cues do drivers interpret other driver's intent to give him priority or not? (and/or the other way around "what tacit or explicit signs do vehicle drivers use to communicate their intent to other drivers?)
	What are the tools different TPs use to interact explicitly and implicitly? (E.g. flashing headlights to inform another driver about yielding the right of way)

[2] We would like to express our very great appreciation to Alexandra Vendelbo-Larsen, Markus Rothmüller and Pernille Holm Rasmussen from Aalborg University for their invaluable assistance in the preparation and conduction of the observation studies.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

"What factors influence interactions"	To what extent do physical variables or social informal dynamics and norms prevail, depending on use case? (For example it is evident that a pedestrian crossing a road with vehicles approaching at 50 km/h, social dynamics might not play the prevalent role)
	Social dynamics of interaction: what -besides local culture- plays a role on who should be given priority depending on the situation? (e.g. pedestrian trying to cross a street during rain / an ambulance flashing or even a car clearly in a hurry.)
	How does the environment influence interaction? What other factors influence interaction?
	What are the influences of demographics (+cultural, +social) on the general participation in traffic and the interaction within?
"What is the threshold for interaction?"	Could there be some kind of Time To Collision or Space Headway derived parameter that can provide a threshold value after which a RU needs to predict the other RU intent through behavioural traits (for example I would imagine a continuum where (i) at very low speeds or at close proximity a lot of explicit gestural signals will be emitted, (ii) at medium distances or medium speeds, bodily behavioural traits would play a significant role on predicting Pedestrians intent, and (iii) at high speeds or from far away most probably physical trajectory parameters will prevail)
	How do TPs behave while not interacting?
	What is the threshold for the need of interaction? How does it depend on surrounding variables (e.g. time of day, age, distance/TTC…)?
Consent and feedback	What types of feedback do HRUs seek from other HRUs once they decide on their trajectory relative to the aforementioned HRUs?
Virtual Testing	Types of consent between HRUs, i.e. “offered priority”, “mutually agreed” or “forced consent”. Apart from individual HRU preference, are there any other variables that may predict the type of consent between interacting HRUs, depending on the scenario?
	What kind of mathematical process model is needed to reproduce observed HRU behaviour in the studied interaction scenarios, in enough detail to allow meaningful simulation-based, virtual testing of AVs?

This report is part of interACT project that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 723395.

3.2 Use cases and chosen real world locations

Following the Deliverable D1.1 “Definition of interACT use cases and scenarios”, locations in Leeds (UK), Athens (Greece) and Munich (Germany) were chosen to observe urban traffic interactions. The must-have use cases were thereby defined as follows:

Table 2: Must-have use cases in interACT (D1.1)

Must-have use cases
Use Case 1: React to crossing of non-motorized TP at crossings without traffic lights
Use Case 2: React to an ambiguous situation at an non-signalized intersection
Use Case 3: React to non-motorized TP at a parking space
Use Case 4: React to vehicles at a parking space

The definitions in Table 2 were intentionally designed to be applicable to a variety of real world locations. As the observation study requires comparability between the different countries to enable the analysis of cross-cultural effects, some limitations were introduced, to control some of the influencing factors (see Chapter 2.2), thus reducing confounding variables.

Limitations regarding the traffic, road layout and infrastructure:

- Intermediately busy roads / intersection / shared spaces (i.e. locations, where the traffic was not fully congested, but road crossings happen regularly)
- One lane in each direction on the main road
- Sidewalk for pedestrians
- Main road has priority and a speed limit of 50 km/h
- Either tall buildings surrounding the observation location or access to CCTV footage (Leeds)

Following these limitations real world locations were chosen to resemble the use cases.

This report is part of interACT project that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 723395.

Use Case 1: Traffic intersection, with one priority road and at least one other road that pedestrians cross regularly, leading to interactions between pedestrians and vehicles.

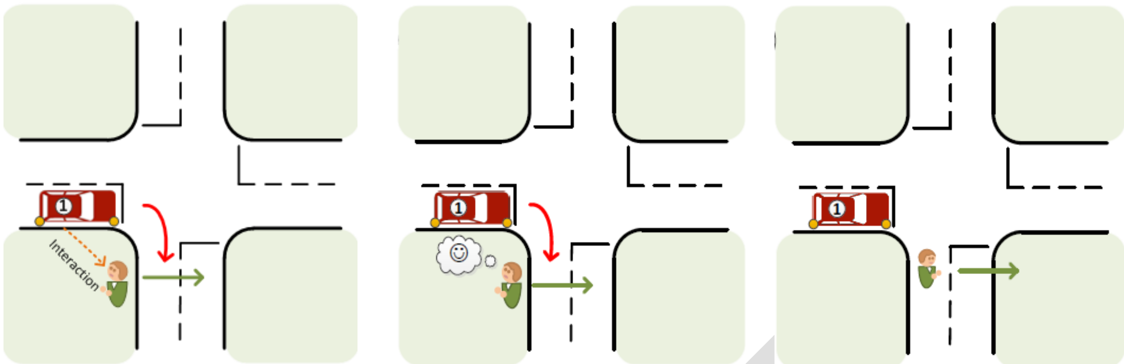


Figure 6: Depiction of a scenario in Use Case 1

Figure 7: Pictures from the locations used for use cases 1 and 2. Top left: Google Maps image from Leeds (UK), top right from Munich (Germany), bottom picture from Athens (Greece)

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Use Case 2: Traffic intersection, with one priority road and at least one other road, where vehicle-to-vehicle encounters happen.

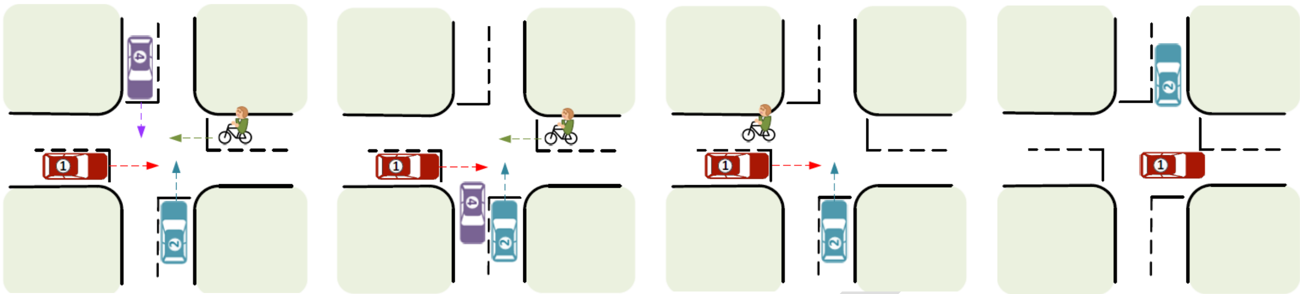


Figure 8: Depiction of a scenario in Use Case 2

While vehicle to vehicle encounters happened quite often at the locations chosen for use case 1 in Leeds and Athens, another location was chosen in Munich to study interactions between drivers.

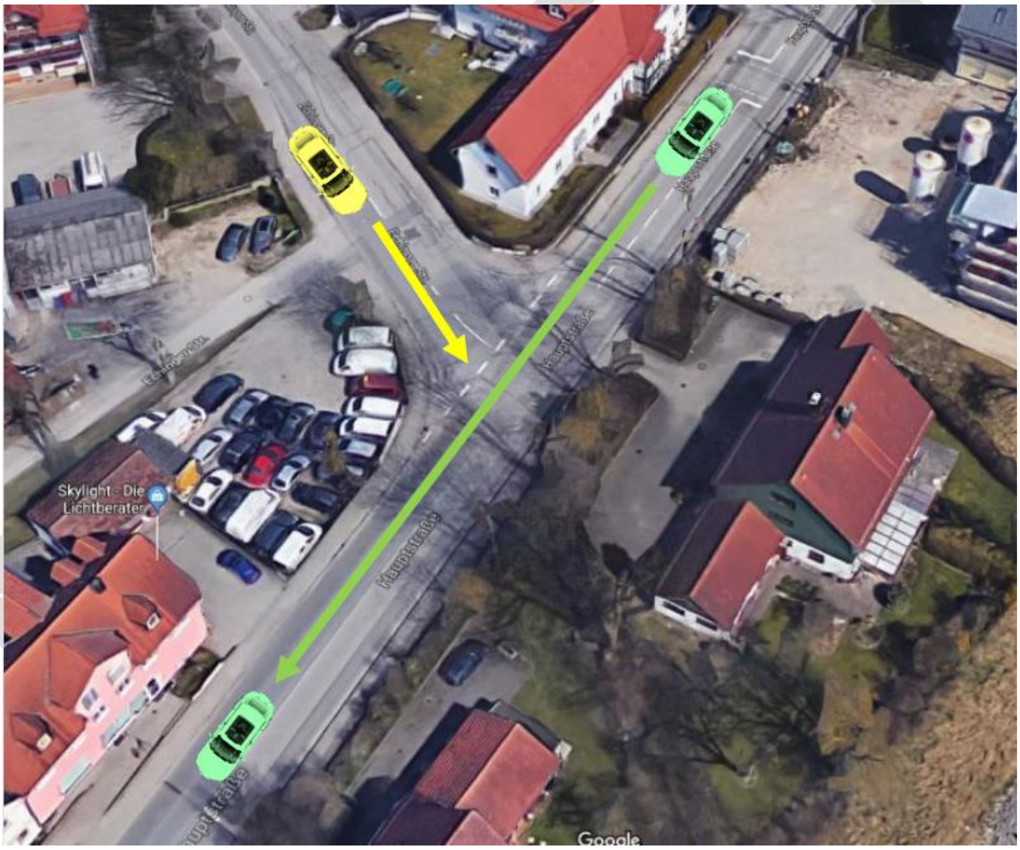


Figure 9: Observed location for use case 2 in Germany

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Use Case 3 and 4: Shared Space with frequent interactions between drivers and other road users, e.g. at a parking lot in front of a point of interest (shopping mall, train station).

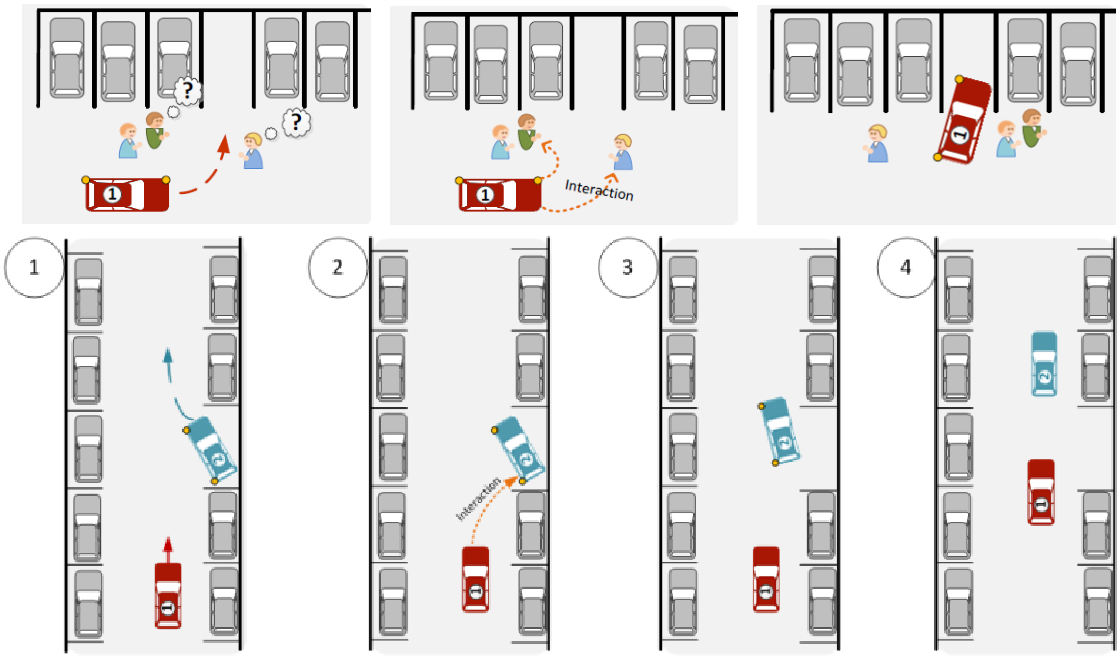


Figure 10: Depiction of a scenario in use case 3 (top) and use case 4 (bottom)

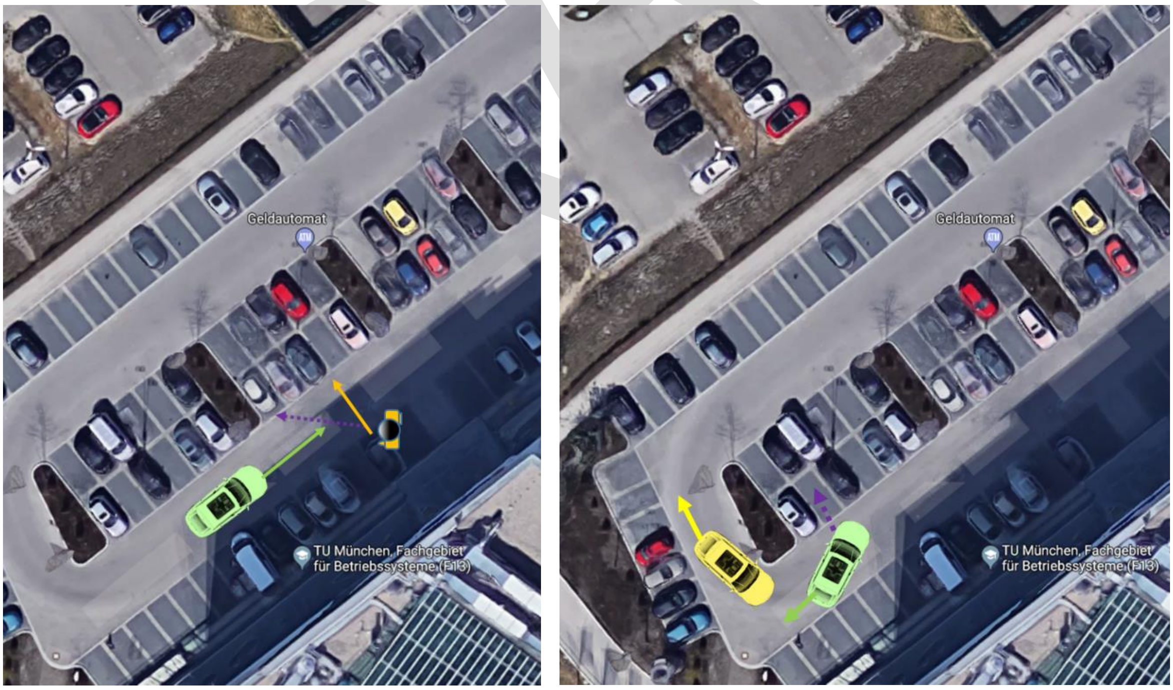


Figure 11: Edited Google images from the locations chosen to observe use case 3 (left) and 4 (right) on a shared space in Germany

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Besides the observation of static locations, the ICCS conducted controlled experiments with drivers participating in unimpeded traffic on a predetermined route in Athens, giving insights into the perception and decision making of drivers in interaction-demanding situations. The driven route for the so called “running commentary”-method can be seen in Figure 12.

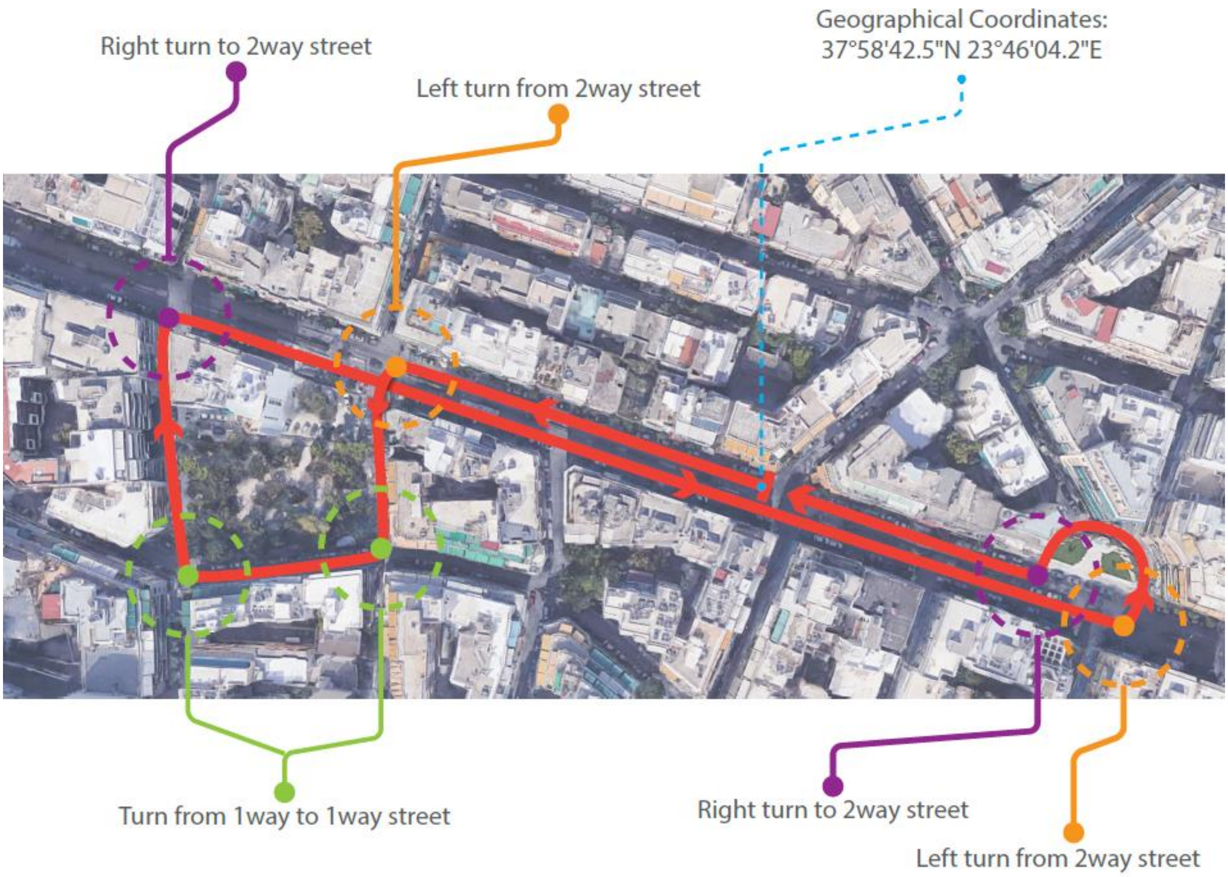


Figure 12: Route for the running commentary observations conducted in Athens (Greece) involving various locations covering all use cases.

This report is part of interACT project that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 723395.

4. Data Collection – Methodology

The traffic observation was conducted utilizing two different study designs to create a holistic understanding about how different HRUs interact in urban traffic.

- **Naturalistic Observation**: at least two experimenters were positioned at a chosen location to manually observe the traffic using a protocol app. Furthermore, logged pedestrians were asked to fill out questionnaires. Where possible a ground based LiDAR and a camera from a higher altitude were utilized.
- **Running Commentary**: controlled experiments with drivers who were asked to drive a specific route. After the experiment (including eye-tracking and video recordings) drivers were asked to comment while reviewing the video of their driving.

Depending on locational circumstances, different methods were deployed. The table below gives an overview of all conducted experiments within the naturalistic observation of T2.1; the utilized methods are described in depth in the subsequent sections of this chapter.

Table 3: Overview of conducted experiments

Research institute	Sub-Study	Utilized methods
ICCS, Athens, Greece	Observation: urban intersection (use case 1 & 2)	Video, observation protocol, questionnaires
ICCS, Athens, Greece	Controlled experiment with drivers (all use cases)	Video from within vehicle, eye tracking, subjective reports by drivers
ITS Leeds, UK	Observation: urban intersection (use case 1 & 2)	Video, observation protocol, questionnaires, LiDAR
ITS Leeds, UK	Observation: shared space (use case 3 & 4)	Video, observation protocol
TUM, Munich, Germany	Observation: urban intersection (use case 1 & 2[3])	Video, observation protocol, questionnaires, LiDAR
TUM, Munich, Germany	Observation: shared space (use case 3 & 4)	Video, observation protocol
TUM, Munich, Germany	Observation: sub-urban intersection (use case 2)	LiDAR, observation protocol

[3] The intersection in Germany had very little vehicle-vehicle interactions. Therefore, the observation for use case 2 was repeated at another location.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

4.1 Questionnaires

4.1.1 Background

The main purpose of the questionnaires was to gain a deeper understanding of the factors which influenced pedestrian decision making when crossing the road in Leeds (UK), Athens (Greece) and Munich (Germany). Based on an evaluation of the previous literature, combined with the research questions defined in section 3, there were three main issues that the questionnaire aimed to address.

Firstly, the questionnaire sought to evaluate whether there were any particular pedestrian characteristics which led to particular actions. For example, research has revealed gender and age differences in road crossing behaviour and accident risk, with many studies showing that female and older pedestrians are more cautious in their street-crossing behaviour than male and younger pedestrians (e.g. Díaz, 2002; Harrell, 1991; Oxley et al., 2005; Rosenbloom et al., 2004). Numerous studies have also shown that pedestrians use cues from other pedestrians to help decide whether or not it is safe to cross an intersection (Hamed, 2001; Marisamynathan & Vedagiri, 2013; Wagner, 1981), and road-crossing wait times decreased as pedestrian flow increased, suggesting that pedestrians are more inclined to cross the road along with others (Zhou et al., 2009). In order to address these issues, questions on pedestrian gender (Q3), age (Q2), driving status (Q14), and the influence of other people (Q11) were all included in the first section of the questionnaire (see Annex 1 for all questionnaire items).

Waiting times for crossing gaps have also been shown to be linked to pedestrian's intended destination (Hamed, 2001), and many of the tactical-level decisions made by pedestrians while planning their journey route are dependent on where they are travelling to and from, and the amount of time available to make the journey (Ishaque & Noland, 2008). For this reason, two questions were included to ask about where participants were travelling from (Q4) and going to (Q5). In addition, the level of familiarity with a particular route has been shown to reduce the amount of time pedestrians are willing to wait to cross, particularly at peak times (Hamed, 2001). The risk taking behaviour of pedestrians has also been linked to their waiting time across a number of studies (e.g. Oxley et al., 2005; Hamed, 2001, Ishaque & Noland, 2008). These issues were addressed by including questions on how regularly participants used the crossing (Q6), and how long they felt they were waiting for a suitable gap (Q7). The intersections for observation were selected based on the fact that there were no clear regulations about who had priority, participants were also asked to evaluate whether they had priority or whether an oncoming vehicle did (Q13). It was anticipated that participants who believed that they had the priority were more likely to engage in risky crossing behaviours.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

A second issue the questionnaire sought to address was whether or not pedestrian characteristics can be used to predict their communication requirements or their interaction styles. For example, Clamann et al. (2017) found that male pedestrians took less time to evaluate their environment prior to making a crossing decision at midblock locations compared to females. For this reason, three questions were included asking about the information from the driver (Q9), the information from the vehicle (Q8), and any other information which pedestrians used to decide when it was safe to cross (Q10). In addition, participants were asked to provide information on any strategies they used to indicate their intention to cross the road, e.g. stepping forward (Q12).

The final issue the questionnaire sought to address was whether there would be any differences in the communication or interaction requirements of more risk-averse pedestrians compared to those who regularly engaged in riskier behaviours. Pedestrians risk-taking propensity was measured using an adapted version of the Adolescent Road User Behaviour Questionnaire (Elliott & Baughan, 2004). The 16-item scale measuring road users' propensity to engage in "unsafe road crossing" was included as question 15 of the questionnaire, and will henceforth be referred to as the RUBQ. All items on this scale were measured using a 5-point Likert scale ranging from "Never" to "Very Often".

The questionnaire provided a subjective measurement of pedestrian's decision making while crossing the road. By administrating the same questionnaire in all three countries, it was hoped that any cross-cultural differences in pedestrian strategies could be captured. In order to establish whether pedestrians' beliefs about their behaviour matched their observed behaviours, all of the questionnaires were linked to the more objective data collected through the corresponding observation protocols (Section 4.2).

4.1.2 Participants

A total of 67, 63 and 52 pedestrians were recruited in Leeds, Athens, and Munich respectively. After a pedestrian was observed in an interactive traffic encounter, one experimenter approached the pedestrian to fill in the questionnaire. Table 4 shows the demographic information of the pedestrians in all three locations. Figures presented pie charts to show where the pedestrians were travelling from and to for Leeds (Figure 13), Athens (Figure 14) and Munich (Figure 15).

Table 4: Demographic information of the pedestrians in all three locations

	Total Pedestrians	Number of Males (%)	Number of Females (%)	Mean Age (SD)	Age Range
Leeds	67	28 (41.79)	38 (58.21)	22.36 (9.90)	16 - 77
Athens	63	49 (77.78)	14 (22.22)	42.37 (14.15)	19 - 74
Munich	52	23 (44.23)	28 (53.85)	35.31 (18.64)	17 - 92

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

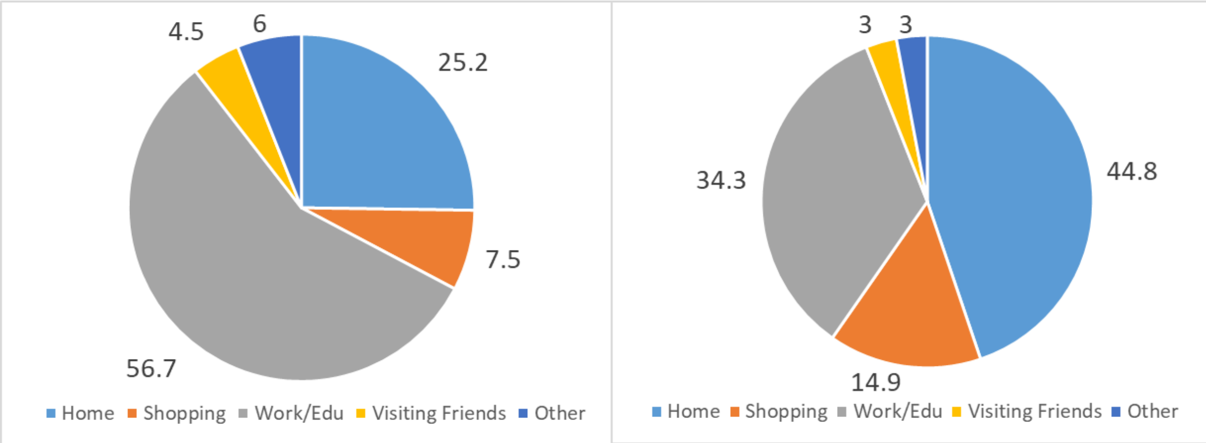


Figure 13: Pie charts (%) where pedestrians were travelling from (left) and to (right) in Leeds.

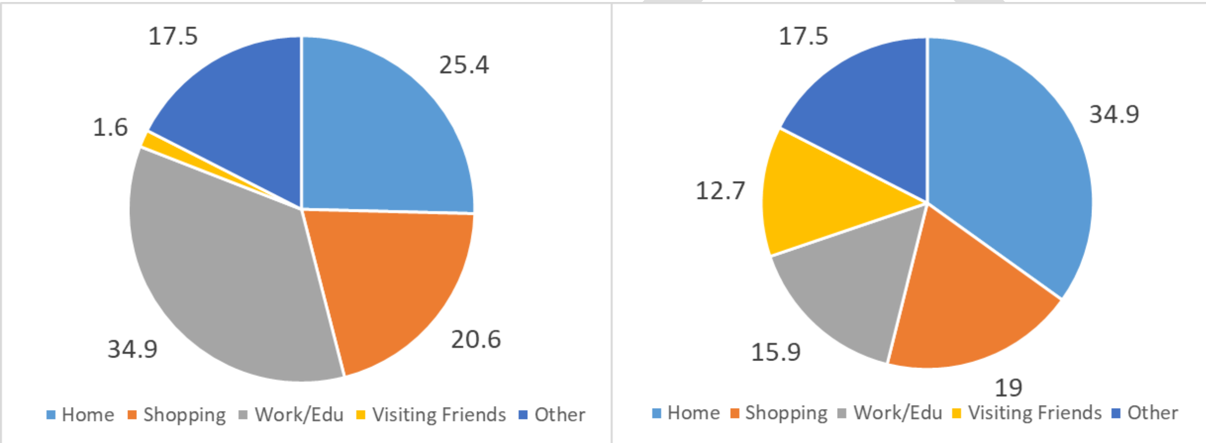


Figure 14: Pie charts (%) where pedestrians were travelling from (left) and to (right) in Athens.

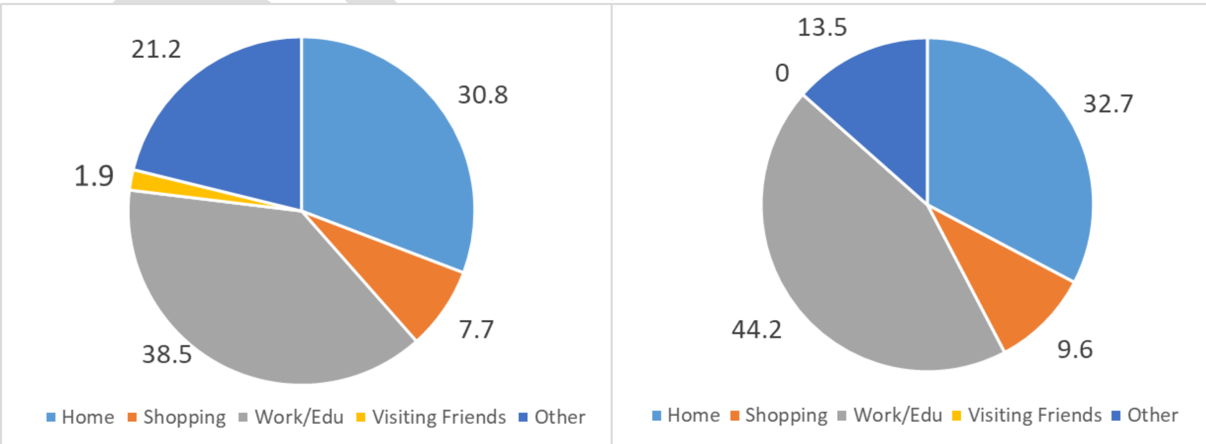


Figure 15: Pie charts (%) where pedestrians were travelling from (left) and to (right) in Munich.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

4.1.3 Stimuli and Procedure

The questionnaire was developed in English (see Annex 1) and then translated into Greek and German by the relevant project partners. Before final administration, each translation was then checked by an independent German or Greek-speaking colleague at Leeds. Ethical consent for the study was obtained from the University of Leeds Ethics Committee. After the questionnaire was developed, responsibility for data collection was assumed by the local representative partners from the University of Leeds, UK; TUM in Munich, Germany; and the Institute of Communications and Computer Systems (ICCS) in Athens, Greece. Data collection was achieved via a Personal Digital Assistant in Leeds and using paper and pencil based questionnaires in Athens and Munich.

The questionnaire administrator worked closely with the observation team to identify suitable participants. Once a particular pedestrian's behaviour had been observed using the Observation Protocol (Section 4.2), the questionnaire administrator approached the pedestrian and asked if they would be willing to complete a short questionnaire. If they responded positively, they were provided with a short verbal introduction to the study and asked to sign a consent form. The questionnaire administrator then read the questions aloud to all participants, and recorded their answers. Participants were not compensated for their time to respond the questionnaires, which took approximately 10 minutes to complete.

4.2 Observation Protocols

4.2.1 Background

The main purpose of the observation protocols was to gain a deeper understanding of the explicit and implicit communication techniques that vehicle drivers and pedestrians used to determine priority at our use case locations. Previous research has identified a number of factors influencing both pedestrian-vehicle interactions and vehicle-vehicle interactions in different settings. Drivers can engage in explicit communication with other road users through the use of eye contact, hand gestures, flashing lights and indicator signals, or implicit communication strategies such as speed reduction (Fuest et al., 2017). Mutual eye-contact has been identified as a factor in facilitating safe interactions between vehicles and VRUs (see Schneemann & Gohl, 2016), with some research suggesting that establishing eye contact with a driver increases the likelihood that the driver will yield to a pedestrian (Guéguen, Meineri, & Eyssartier, 2015). At greater distances, drivers are more likely to use implicit communication strategies to convey their intent. For example, interview data collected by Sûcha (2014) showed that drivers make use of a variety of techniques to force pedestrians to yield, including refusing to decelerate, speeding up, and driving more in the centre of the road to avoid hitting a pedestrian while not stopping for them. Finally, physical factors such as traffic volume (Harrell, 1991; Hagel et al., 2014), darkness and weather conditions (Klop and Khatak, 2007; Sayed et al., 2013), are also likely to affect crossing behaviour. Finally, pedestrian attention is also likely to impact on their crossing behaviours. Hatfield and Murphy (2007) investigated the effect of mobile phone use on pedestrian crossing behaviour by comparing different groups of pedestrians, and found that pedestrians who crossed while talking on a mobile phone crossed more slowly and were less

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

likely to look at traffic before starting to cross. In a simulator study, Schwebel et al. (2012) found no safety effects of holding a mobile phone conversation, but showed that listening to music or texting led to more unsafe crossing decisions. Thus, it is likely that AVs will need to be able to recognise distracted pedestrians and adjust their driving behaviour accordingly to avoid risky situations occurring.

The observation protocols were developed to capture detailed information on explicit and implicit cues used by drivers and pedestrians that may not be visible through overhead camera recordings. It was also anticipated that by having trained field-workers observe a subset of the recorded interactions in real time, this would improve the quality of tracking possible within the video recordings by enabling a more detailed coding of interactions. Three types of protocol were developed, with the first focusing on pedestrian – vehicle interactions (Use Case 1), the second protocol focusing on vehicle – vehicle interactions (Use Case 2 and 4) and the third focussing on pedestrian – vehicle interactions on shared spaces (Use Case 3).

The pedestrian- vehicle observation protocol was divided into 6 main sections (see Annex):

- An approaching phase divided into “Pedestrian Analysis” and “Vehicle & Driver Analysis” sections, where the behaviour of both the vehicle and pedestrian was monitored as they approached the intersection. Information for this analysis included traffic participants’ speed, head movements, gestures, and any distracting activities prior to the point at which the pedestrian reached the edge of the road
- A crossing phase, divided into “Pedestrian Analysis” and “Vehicle & Driver Analysis” sections, where the behaviour of both the vehicle and pedestrian was monitored from the point at which the pedestrian reached the edge of the road. Information captured included whether or not the vehicle or the pedestrian stopped, any explicit communication they engaged in (e.g. hand gestures, head movement, flashing lights), and any implicit cues that were provided by both partied (e.g. pedestrian stepping out onto the road / vehicle decelerating)
- A general information section where the pedestrian’s demographic information, the level of traffic flow, the weather conditions, and the time of day were recorded
- A schematic representation of the junction where the observers could provide a drawing of the direction in which all of the observed traffic participants moved.

The vehicle – vehicle interactions were divided into 4 main sections:

- Vehicle 1 Analysis, where the actions of the vehicle showing intent (e.g. turning) was captured. This included information on the vehicle movements, signals used, and any hand gestures or head movements observed.
- Vehicle 2 Analysis, where the actions of the interacting vehicle were captured. Once again, this included information on the vehicle movements, signals used, and any hand gestures or head movements observed.
- A general information section where the vehicle description, level of traffic flow, the weather conditions, and the time of day were recorded.

This report is part of interACT project that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 723395.

- A schematic representation of the junction where the observers could provide a drawing of the direction in which all of the observed traffic participants moved.

4.2.2 Pedestrian & Vehicle Selection

The observation teams consisted of at least two people, who shared the observation and questionnaire administration roles. The observers were situated at a location far enough away from the intersection to avoid influencing the behaviour of road users, but close enough to observe the

important details of any interaction (see Location X & Y in Figure 16).

Figure 16: Positioning of observers for use cases 1 and 2 with left-hand traffic in Leeds, UK.

Pedestrian – Vehicle interactions

For the pedestrian – vehicle interactions, one of the researchers would be responsible for choosing a pedestrian to observe. The pilot research had shown that choosing a pedestrian together as a whole group proved to be quite difficult as interactions emerged and passed quickly, so before agreeing on a pedestrian, it was already too late. The pedestrians were selected to have a broad representation of both gender and age categories, from children to older age groups. The choosing of a suitable pedestrian happened approximately 5-10 meters before the pedestrian entered the road, meaning that the observation started as a pedestrian approached the intersection. This resulted in some pedestrian observations which did not contain any explicit interaction with approaching vehicles, but instead provided information on gap acceptance and pedestrian searching techniques.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Two of the observers would monitor each interaction, with one observer focusing on each interaction participant i.e. one observer monitoring pedestrian behaviour and one monitoring vehicle behaviour. In order to accurately capture the sequence in which the interplay of the interacting parties took place, the observers discussed aloud the pedestrian and driver/vehicle actions throughout the interaction. Once the interaction was completed, the sequencing order was once again discussed to ensure that it was captured as accurately as possible. This technique was practiced extensively prior to starting the main data collection process, and where possible the questionnaire administrator also observed the interaction to provide further confirmation that the correct order had been captured.

Vehicle – Vehicle interactions

As with the pedestrian – vehicle interactions, one of the observers chose a vehicle for the observation. This decision was made before any interaction occurred, and any signalling or change in trajectory happened. One observer then focused on the vehicle that was chosen, while the other focused on the vehicle it interacted with, if there is one. Not all observations involved direct communication between two vehicles. For example, a vehicle decelerating and indicating to turn left without any direct interaction with other vehicles close by but some in the far distance, was still counted as an interaction. Observations were only disregarded if there was clearly no other vehicle on the road that might influence the first vehicle's actions.

Two of the observers would monitor each interaction, with one observer focusing on each interaction participant i.e. one observer monitoring vehicle 1 and one monitoring vehicle 2. In order to accurately capture the sequence in which the interplay of the interacting parties took place, the observers discussed aloud the pedestrian and driver/vehicle actions throughout the interaction. Once the interaction was completed, the sequencing order was once again discussed to ensure that it was captured as accurately as possible. This technique was practiced extensively prior to starting the main data collection process, and where possible the questionnaire administrator also observed the interaction to provide further confirmation that the correct order had been captured.

4.2.3 HTML Based App

At first, the protocols were developed in Microsoft (MS) Excel and tested by using a pen on printouts. To simplify the data preparation from the observation protocols, enable measurements synchronized in time[4] and reduce the amount of paper used within the observation, the protocols were transferred into an app that was programmed at the TUM and usable on a variety on smartphones and tablets. The following prerequisites were formulated for the app in the development phase:

- Usable on different devices, running different operating systems
- Utilize all elements of the observation protocol (including free text inputs or drawings) while being usable (i.e. big enough buttons)

[4] The app ensured that the videos, LiDAR and observation app all had the same time stamp, thus, manually observed interactions could easily be found in the LiDAR data and videos.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

- Displaying a timestamp with the device's system time to synchronize the video recordings to the observation protocols as well as the questionnaires
- Capture the time of the start of an observation
- No server dependency / run offline
- Save data locally as a .csv file

As different devices (involving different operating systems and screen sizes) were used, the app was programmed using HTML5, CSS, JavaScript and the jQuery JavaScript library.

The app resembled the Excel based protocols and stays usable by browse buttons that navigate through the different observation phases. Each button press was recorded with a timestamp and a sequence number, which was displayed on the button. This enabled multiple button presses within one encounter – e.g. if a pedestrian turned his head left right left.

To depict the observed encounter/interaction, a bird's eye view picture was loaded onto the screen. The user could then drop icons of pedestrians and vehicles onto the scenery and draw lines and arrows by tapping once for the starting point and once for the ending point. All clicks on the depiction were recorded so that the depictions could be recreated from the csv file and the picture of the scenery.

A dedicated sync button displays the device's UNIX time for ten seconds in 200ms steps. By showing the device to a camera, the video recordings could be synchronized with the app's observation files. A back button is implemented to remove false inputs. After finishing the typing on the tablet, the observation can be exported to be locally saved as a csv file.

Three Apps were programmed, to observe different encounters:

- Pedestrian – Vehicle Interaction (for use case 1)
- Vehicle –Vehicle Interaction (for use cases 2 and 4)
- Shared Space: Pedestrian – Vehicle Interaction (for use case 3)

The MS Excel based observation protocols and pictures of the pedestrian-vehicle observation app can be found in Annex 2.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

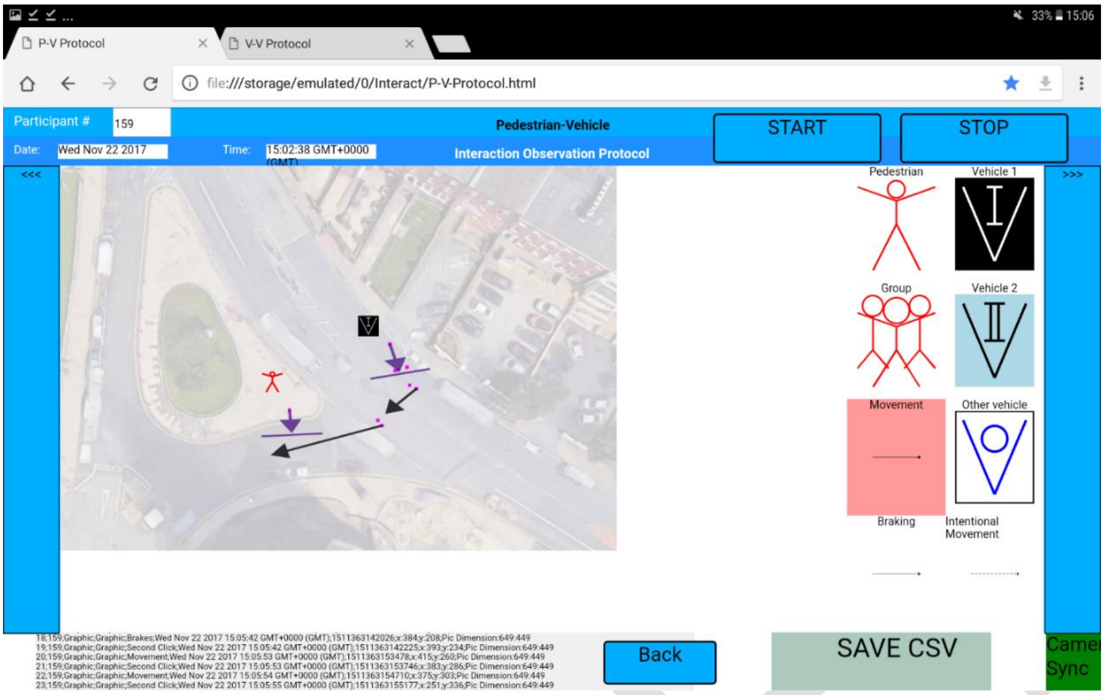


Figure 17: Screenshot of the depiction tab of the interACT observation app for the pedestrian to vehicle interaction (P-V-Protocol.html)

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

4.3 Video

The video recordings were made in accordance to the individual national data privacy policies. In Athens and Munich, GoPro cameras were placed on high ground (e.g. 4th floor of a building) to record the point of interest. This procedure ensured, that recorded persons and number plates are not identifiable from the videos, due to the angle and the low pixel count.

In Leeds, an outdoor HD wireless IP camera was mounted on the roof of the Laidlaw library. The camera is composed of a colour sensor CMOS, a wireless antenna, and an Infrared lamp array. The camera can automatically switch to Infrared recording mode by night such that we could have a high quality video by day and night. The camera can be configured and communicate via TCP/IP protocol with any device computer, phone, tablet etc. The camera settings were the following:

- Camera model: Foscam FI9803P
- Codec: H264-MPEG-4 AVC(part 10) (h264)
- Resolution: 1920 x 1080
- Display Resolution: 1920 x 1080
- Format: Planar 4:2:0 YUV

Figure 18: Camera used for the observations

The camera was connected to a Wi-Fi router which was communicating with a laptop, both placed inside a waterproof box. The laptop was running Ubuntu 16.04. During the observations, the protocol was to have one person going to the roof and use a personal laptop to connect via SSH to the distant laptop left in the waterproof box in order to activate the camera. Once the observations were finished, we either deactivate the camera or leave it running for several days.

We implemented a Python script that was able to create the connection with the camera, record the scene and save a new video file every hour. It was possible to check if the recording were being done correctly by using software such as WinSCP which allows connecting to the distant laptop and access to the video files directory.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

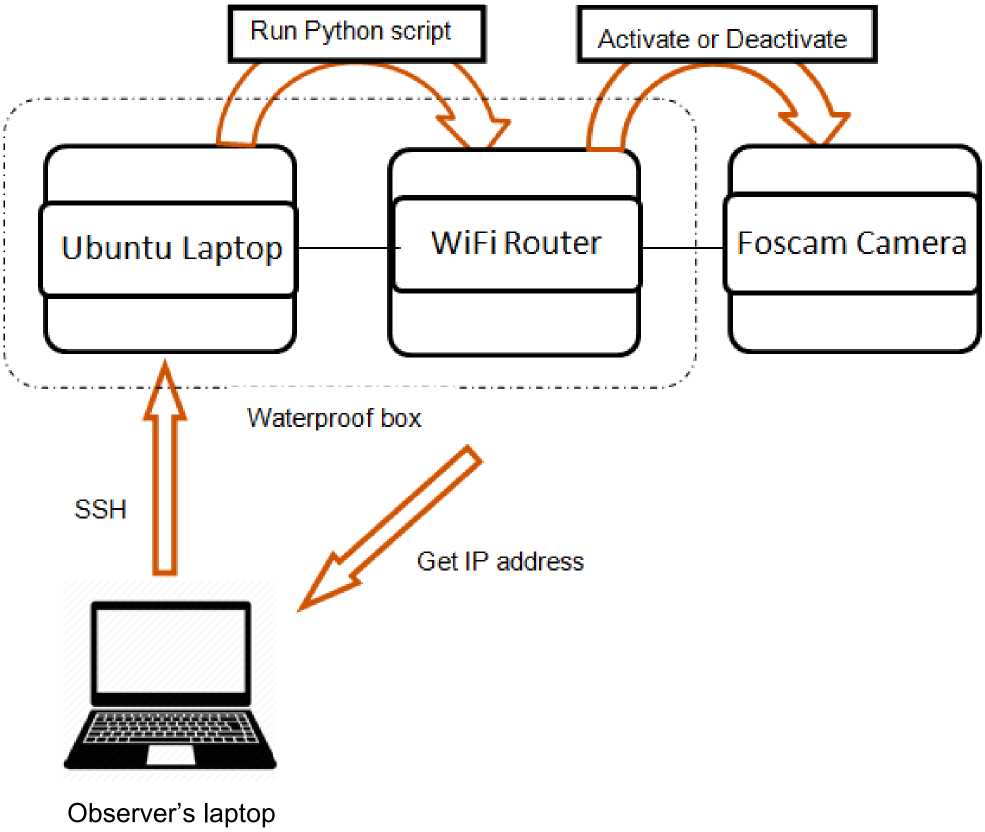


Figure 19: Camera video recording protocol

With about one complete month of video recording, we obtained about 600 hours of video data containing vehicle-vehicle interactions and vehicle-pedestrian interactions in the Woodhouse Lane intersection in Leeds. To better understand these interactions, we will make use of these videos for pedestrian and vehicle tracking purposes.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

4.4 LiDAR

The main purpose of the LiDAR observation was to receive synchronized quantitative measurements of the position, velocity and the type of traffic participants in addition to the questionnaires, observation protocols and videos. Compared with video observation, the LiDAR observations were made with the same sensors like in the CRF experimental vehicle and a similar perspective. The LiDAR gives a more accurate position and velocity estimation of traffic participants compared to the video observation.

Table 5 gives an overview of the most important technical facts of the used LiDAR sensor. Furthermore, the sensor provides an object tracking with object properties position, size-and velocity of traffic objects.

Table 5: ibeo LUX HD - Technical facts

Ibeo LUX HD	
Range	90m @ 90% remission 30m @ 10% remission
Horizontal field of view	2 layers: 110 deg (50 deg to -60 deg) 4 layers: 85 deg (35 deg to -50 deg)
Vertical field of view	3.2 deg
Data update rate	12.5 Hz
Accuracy (distance independent)	10cm
Angular resolution	Horizontal: up to 0.25 degree Vertical: 0.8 degree
Distance Resolution	4cm

For an easy use, the ibeo LUX LiDAR sensor was integrated in a housing with power supply, a hard disk storage and a GNSS receiver.

4.4.1 System structure of the LiDAR observation box

The LiDAR observation box consists of a power supply, which provides enough energy to drive the whole box for about twelve hours, a hard disk drive to record measurements, a GNSS receiver that gives an accurate UTC timestamp, an ibeo LUX LiDAR and a Raspberry Pi as signal processor. Furthermore, an optional webcam can be used to provide a video image during the installation phase. Figure 20 gives an overview of the main components inside the observation box.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

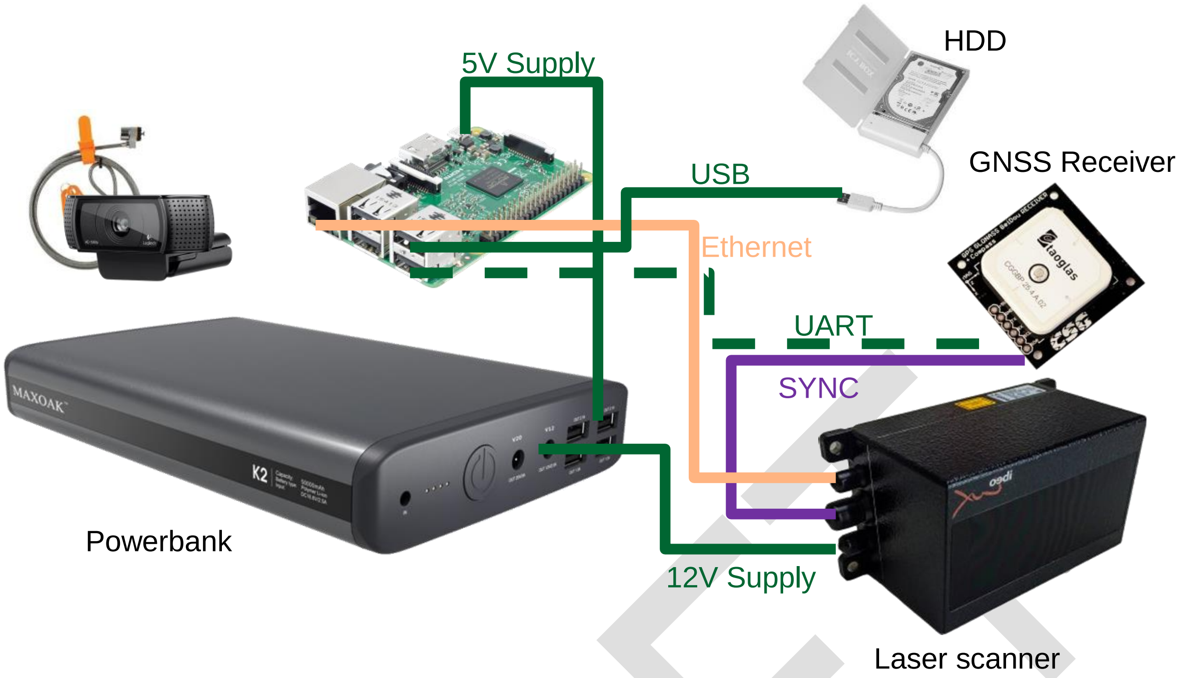


Figure 20: System structure of the ibeo LUX observation box

To synchronize the LiDAR measurements with video observation and observation protocols, the accurate UTC timestamp provided by GNSS receiver was also recorded on the HDD.

Figure 21: Depiction of the finalized prototype used by Observers in Leeds

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

4.5 Interactions from the point of view of the driver

In addition to the static observation, the ICCS ran an experiment in Athens focussing on drivers. The aim of the study was to observe interactions between drivers and between drivers and pedestrians in urban environment from the point of view of a driver and understand how such interactions evolve.

Twenty-one experienced drivers were asked to drive their own passenger car in a predefined course, while wearing an eye glass mounted gaze sensor. There were 10 male and 11 female drivers, their mean age was 39.1 years (median 38 years, standard deviation 11.7 years) and their mean driving experience was 18.5 years.

The course consisted of a circular route of 0.75 km which was driven 5 times by each subject. The total course length was 3.75 km and the mean driving duration was 18 minutes. The course included left turning from a two-way street without a traffic light, right turning from a smaller to a two-way street, straight segments where pedestrians frequently cross and small one-way streets where pedestrians frequently walk on the street. It was expected that there would be a lot of interactions between drivers relevant to the left and right turns of the participants and a lot of interactions of the participants with pedestrians who would wish to cross the street or who would walk on the street.

After the end of each driving session, the participant (subject) was asked to watch selected parts of the eye gaze video recording and to comment aloud on the process of his/her decision making for each case of interaction with another driver or with a pedestrian. Verbal protocols offer a way to record the human thought process (Ericsson & Simon, 1993) and have been used in driving studies (Portouli et al, 2014).

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

5. Data analysis

5.1 Questionnaires

Data collected from Leeds, Munich and Athens was transformed into the same template in an excel sheet to ensure consistency for analysis. A series of data were analysed separately for each location by University of Leeds using SPSS and reported in Section 5.3.1.

First, in order to understand how many pedestrians have reported that they use certain vehicle information and driver information to decide crossing, we have provided the descriptive data for each of the information used. The same was conducted for how pedestrians indicate their intention to cross. The average used for each of these measures (vehicle information, driver information and indicating intention) was calculated and was compared to 0 (0 indicating not using the information) by using one-sample t-tests respectively. This is to investigate whether pedestrians have been using the information or not. Paired-sample t-tests were also used to compare the average of vehicle information used and driver information used to investigate whether pedestrians are more likely to use one of these information to decide crossing.

Second, the RUBQ was investigated by providing the proportion of pedestrians who responded 'Never', 'Rarely', 'Sometimes', 'Often' and 'Very Often' for each of the questions. An average RUBQ score was also calculated for each pedestrians.

Third, the effect of gender, who pedestrians think has the priority ('you' or the driver), driver (whether they are a driver or not), and the effect of how other pedestrians affecting their crossing behaviour were tested to investigate whether each of these measures has an effect on one another. The effect of each of these measures (gender, priority, driver and people effect) on ratings (e.g. familiarity, safety, gap length, average vehicle information used, average driver information used, average intention cues indicated, average RUBQ scores) were also investigated .

Forth, correlations between ratings were conducted to investigate the relationships between them.

5.2 Observation Protocols

The Observation Protocols were saved individually for each observed encounter. Python scripts were developed to merge the observation protocols of one use case and location into one MS Excel file. Within the file, each row represents one encounter and each column represents one possible button press. The sequences of the button presses were put into the appropriate columns and enable to research patterns within individual scenarios. For deliverable 2.1, frequencies of occurrences were analysed to evaluate comparisons and differences and to enable the creation of sequence diagrams for different encounters (see chapter 7.1). The observation protocols will be further analysed within Task 2.2 to research observed sequences in depth to model road user behaviour.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

5.3 Analysis of data from the running commentary study

An analyst watched the subject's eye gaze and scene video as well as his/her retrospective commentary, focusing on any occurrence of interaction between the study participant and another driver and between the subject and a pedestrian.

Interactions among drivers were analysed relevant to the left turn from a two-way street and right turn to a two-way street in the locations shown in Figure 22. The interaction start was set as the time point when i) the subject had to wait for a gap in the oncoming traffic before turning or ii) the subject started turning knowing that the oncoming driver would have to modify his/her vehicle motion. For each interaction, the analyst annotated the type of the interacting vehicle and whether the other driver reacted. The sequence of signals or cues by the subject and his/her vehicle and by the other driver and his/her vehicle was annotated for each interaction. The subject's commentary was transcribed in digital format.

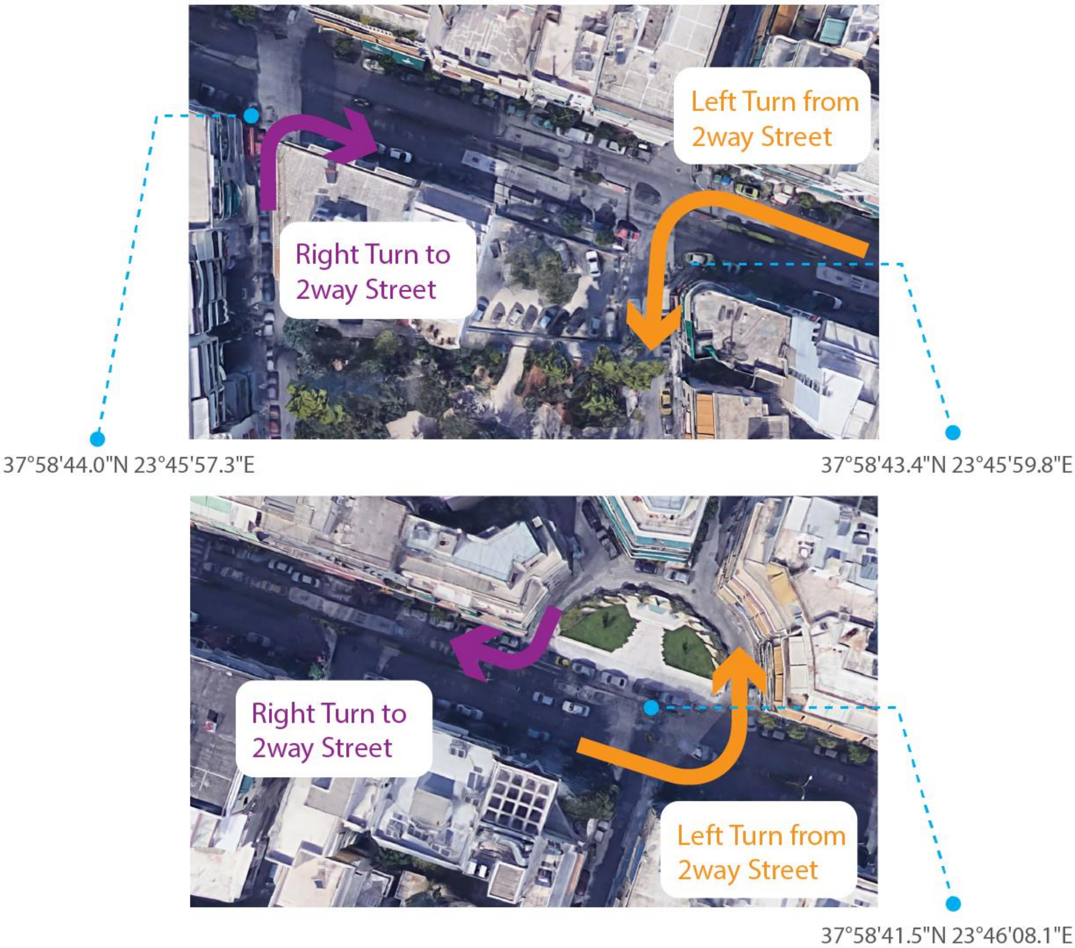


Figure 22: Locations where interactions among drivers were analysed

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Figure 23: Example of eye gaze video recording during a left turn from two-way street with oncoming traffic

Figure 24: Example of eye gaze video recording during right turn to two-way street

Interactions between participating drivers and pedestrians were analysed both on intersections and on straight road segments. An encounter with a pedestrian was considered as interaction when the pedestrian in the vicinity of the participant driver (i) affected the car movement and/or the driver's behaviour in an observable manner and (ii) was the object of at least one eye-fixation from the driver. The driver-pedestrian interactions were categorized according to the pedestrian's orientation to the street axis into:

a) Crossing interactions i.e. with pedestrians' orientation perpendicular to the street axis (with probable intention to cross) and
b) Parallel interactions i.e. with pedestrians walking on the street roughly following the street axis, in the same direction or opposite to the subject's vehicle (with a probable intention to share road space).

For pedestrians with perpendicular orientation to the road, the interaction start was set as the time point when there was a first cue by the pedestrian interpreted by the driver as intention to cross. For pedestrians with orientation parallel to the road the interaction start was set at the time point of the first eye fixation of the driver on the pedestrian.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

For each interaction, movement behaviour, cues and signals were annotated for pedestrians. For the participating drivers, the vehicle movement, signals and fixations towards the pedestrian were annotated along with the spontaneous driver's utterances relevant to each interaction. In addition, the participants' video-assisted retrospective commentary was transcribed in digital format.

Figure 25: Example of eye gaze video recording for a pedestrian wishing to cross

Figure 26: Example of eye gaze video recording for a pedestrian walking on the street

5.4 Video

To identify communication processes in traffic and extract kinematic information of all road users involved in an interaction, we applied some tracking algorithms of both vehicles and pedestrians on the video data collected during the observations. Tracking is a challenge for computer vision systems because of the multiple uncertainties (e.g. occlusions etc.) due to complex environments. Multiple Object Tracking (MOT) requires first to determine the number of objects to track and second to be able to identify each object as such MOT can be seen as a multi-variable estimation problem, where the goal is to perform a MAP (Maximum A Posteriori) estimation of the sequential states of the objects (Luo, Zhao and Kim, 2014). Tracking can be easier by using a fixed camera in the environment and make use of temporal information from a sequence of frames.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

We first applied a blob tracking of pedestrians and vehicles on the videos. Then we used machine learning and deep learning approaches to vehicle detection for comparison purposes and to give us an idea of the kind of the technique that we will need to use for better tracking results.

5.4.1 Vehicle and Pedestrian Tracking using OpenCV Blob tracking

A blob can be defined as a group of connected pixels. First thresholding converts images into binary images, then connected components are extracted by finding the contours and centres. Finally connected pixels and close centres are grouped as blobs (OpenCV.org).

Before blob tracking, pedestrians and vehicles are detected by subtracting the image background for each video frame. Background subtraction builds a background model used as a reference model in order to detect moving objects (Gouda, 2015). Background Modelling is based on the assumption that the background is static. It consists in extracting an estimate of the background from the rest of the image by using some methods such as mean filter, a running Gaussian average etc. There are two variants: recursive algorithm which updates each frame the estimate of the background and non-recursive algorithm which stores a buffer with the previous frames and the background estimate from them.

Once the background is extracted, the remaining foreground mask is passed to the blob tracking algorithm which computes the centroid, the **ID** and the angle of the moving objects. A bounding box is usually drawn around the object.

We used an open source C++ code (Sobral, 2014) which combines background subtraction and blob tracking for pedestrians and vehicles. For each frame, the first step consists in removing the image background using BGSLibrary a background subtraction library (Sobral, 2013) based on OpenCV, computer vision library.

System settings:

- Ubuntu 16.04
- OpenCV 3.3

Observations:

The blob tracking software provide satisfying results, pedestrians, cyclists and vehicles are correctly tracked. A white dot is displayed on the images to show the trajectory of the moving objects (cf. output images) and the main orientation of the body is represented by a green line. The blob tracking works well both during the day and at the night.

However, some improvements are needed as the tracked ID changes sometimes over time even if it is the same blob. The blob tracking is not distinguishing a pedestrian from a vehicle, we need to add a classification method either based on the speed or on the shape of the blob.

The next step will consist in storing the tracking information such the ID, the speed, trajectory etc. inside a SQL database. This database will allow us to perform some statistics analysis on the videos.

Some output images from the blob tracking are displayed below.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Figure 27: Example images of blob tracking in daytime at the intersection in Leeds

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Figure 28: Example images of blob tracking in night time at the intersection in Leeds

5.4.2 Vehicle Detection using HOG + SVM classifier

A commonly used detection technique is to combine HOG (Histogram of Oriented Gradients) with SVM (Support Vector Machines). HOG is a technique that was invented for the purpose of human detection (Dalal and Triggs, 2005). SVM is a binary classifier trained with the HOG descriptor of a vehicle as a positive label. After training, the obtained classifier is then able to determine whether a proposed HOG is that of a vehicle or not. OpenCV has already an implementation of an HOG and SVM detector.

We used the HOG + SVM open source vehicle detector implemented by (Özlu, 2017).

System settings:

- Ubuntu 16.04

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

- OpenCV 3.3

Figure 29: Example image of vehicle detection using HOG + SVM Classifier

Observations:

The HOG + SVM vehicle detector is not very efficient; as it has a lot of false detections which will make the tracking and data association much harder. British buses are not well detected with this method as the training image set didn't contain this kind of vehicle.

5.4.3 Vehicle Detection using Deep Learning

Deep neural network is currently the most popular technique to perform object detection. As an unsupervised method, it doesn't require labelling the object but the neural networks have to be trained with a lot of data in order to be efficient. Keras and TensorFlow are the open source libraries used for the training.

We used a deep learning open source library (Özlu, 2018) to detect vehicles in our videos. This source code provides a model that has been already trained. The detection works very well when the vehicles are not occluded.

System settings:

- Ubuntu 16.04
- Python 3.3
- Keras 2.2
- TensorFlow
- OpenCV 3.3

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Figure 30: Example images using deep learning detection and classification

Observations:

The detection using deep neural networks works very well a part from some occlusion problems. We are planning to use a deep learning based vehicle detector combined with an open source pedestrian and vehicle tracker (Hanheide and Dondrup, 2014) in order to get better tracking results and provide a better analysis and understanding of vehicle-vehicle and vehicle-pedestrian interactions.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

6. Results

6.1 Interaction Vocabulary

The following section describes the types of signs vehicle drivers use to communicate with other road users derived from the running commentary experiment conducted by ICCS in Athens. Different types of signs were exchanged between:

- Driver of passenger car – Another driver of passenger car
- Driver of passenger car – Pedestrian (from the driver's point of view)

Table 6 describes the types of signs as identified and classified from the eye-tracking videos and the retrospective think-aloud sessions by participant drivers in the ICCS driver observation study. **These types of signs primarily reflect the drivers' point of view and are still subject to review.**

The signs are classified below only in terms of their observable, objective manifestation (e.g. horn, turn indicator, head nodding, body/head orientation) without the possible meaning that each one has/or might take. The observable manifestation of the signs is hereafter referred to as **"physical signifier"**. Allocation of specific meaning to these physical signifiers (i.e. the signified aspect) is a task to be done at a second phase once the types of physical signifiers are stabilized.

Also, in the classification below, no distinction is made between explicitly emitted and implicit/unintended types of signs (although some types e.g. turn-indicator are by definition explicitly emitted, others e.g. body/head orientation, are always subject to interpretation even by the performing subject itself).

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Table 6: Interaction Vocabulary

Traffic Participant	type of physical signifier	Sign type (physical signifier)	examples
Car/Driver	Driver's behaviour	**Hand gesture**	e.g. move hand sideways, show palm
		Head Nodding	e.g. sideways, downwards, …
		Eye-contact	e.g. with pedestrian, with other driver
	Car	**Car movement**	e.g. accelerate, kept pace, Stopped, turned
		Car positioning	e.g. protruding on intersection, keeps left/right
		Engine noise	e.g. rev-up the engine on idle
	Car HMI	**Turn indicator**	Left / right
		Headlights flashing	
		Horn	e.g. one long press, one momentary, two….
		Alarm indicator	
Pedestrian	Pedestrian's body	**Hand gesture**	e.g. raised hand, extending palm, waving …
		Head Nodding	e.g. sideways, downwards, …
		Eye-contact (with car driver)	
		Gaze towards car (when it is clear that the pedestrian has seen the car)	
		Head/body orientation (combined since semantically they form a whole)	e.g. facing car, facing sideways, …
		Body movement	e.g. walking parallel towards car, hesitating, accelerating, …

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

It should be noted that further full codification of each "Sign Type" to more detailed physical signifiers as presented in the "examples" column, tends to become impractical. This is because these signifiers are either contextually dependent (e.g. Head/body orientation can be meaningfully specified only by considering the particular situation and manoeuvre) or they are non-exhaustive (e.g. Hand gesture – *raised hand in front / raised hand sideward / waved … / extend palm etc.* – can never be fully objectively codified).

A more detailed specification of "Sign Type" can be useful only if it also includes the "meaning" (i.e. the signified part) of the sign (e.g. waved hand so as to say thanks, or nodded to signify that he lets somebody pass). However, for the purposes of interACT in the analysis phase it is of value to distinguish between the objective and the semantic aspects of the signs, which will be derived in the future and described in Deliverable 2.2 "Final description of psychological models on human-human and human-automation interaction".

6.2 Questionnaires

Vehicle Information Used - In the questionnaire, we asked pedestrians what information from the vehicle they used to decide whether it was safe to cross, and they were allowed to choose more than one options. Figures show visual representations of the percentages (%) of pedestrians who said 'no' (red bars) and 'yes' (green bars) as well as in number of pedestrians who said 'no' written on each bars for Leeds (top), Athens (middle) and Munich (bottom) (Figure 31) respectively. Table 7 provides information on pedestrians (numbers and %) who responded yes in each location as a comparison.

Table 7: Number and percentages of pedestrians who responded 'yes' for each vehicle information from each location

	Among 67 pedestrians from Leeds		Among 63 pedestrians from Athens		Among 52 pedestrians from Munich	
Vehicle Information	Yes (pedestrians)	Yes (%)	Yes (pedestrians)	Yes (%)	Yes (pedestrians)	Yes (%)
Speed	40	59.70	5	7.94	30	57.70
Distance	35	52.24	15	23.81	32	61.54
Braking	14	20.90	11	17.46	6	11.54
Flashing	12	17.91	0	0.00	1	1.92
Trajectory	11	16.42	7	11.11	5	9.62
Indicator	28	41.79	0	0.00	2	3.85
Passing	21	31.34	26	41.27	17	40.38
None	21	31.34	9	14.29	1	1.92

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

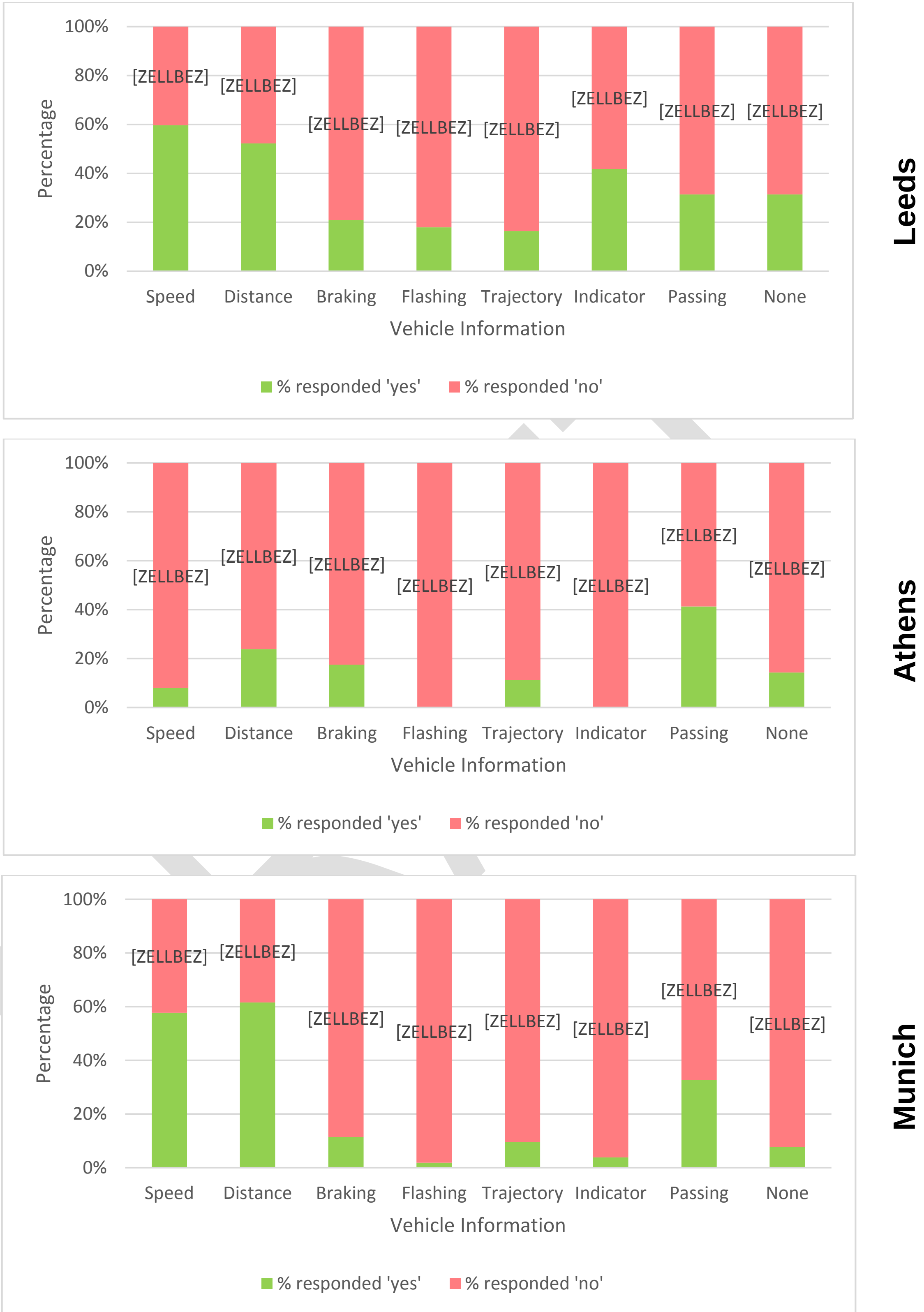


Figure 31: Percentages (%) of pedestrians reported 'yes' (green) and 'no' (red) for each of the vehicle information used in Leeds (top), Athens (middle) and Munich (bottom). Numbers written on the red bars indicating the number of pedestrians who responded 'no' in each.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Among 67 pedestrians in Leeds, 46 reported that they use vehicle information (68.66%), where the top three mostly used vehicle information reported are the 'speed', followed by 'distance' and 'indicator'. In Athens, 54 out of 63 reported that they use vehicle information (85.71%) and the three mostly used vehicle information reported are 'passing', followed by 'distance' and 'braking'; whereas in Munich, 48 out of 52 pedestrians reported that they do use vehicle information (92.31%) and the top three used information are 'distance', 'speed' and 'passing'.

Average of vehicle information used by each pedestrian were calculated by taking the sum of scores for each vehicle information used and divided by 7 (speed, distance, braking, flashing, trajectory, indicator, and passing). One-sample t-tests revealed that the average of vehicle information used by pedestrians in Leeds was significantly higher than 0, $t(66) = 17.63$, $p < .001$; same for Athens, $t(62) = 13.24$, $p < .001$ and Munich, $t(51) = 10.96$, $p < .001$.

Driver Information Used - We also asked what information from the driver did pedestrian use to decide whether it was safe to cross (see Table 8 for data from all locations) and they were allow to choose more than one from the options. Figure 32 provides visual representation of each driver information used by pedestrians in Leeds, Athens and Munich respectively.

Table 8: Number and percentages of pedestrians who responded 'yes' for each vehicle information from each location

	Among 67 pedestrians from Leeds		Among 63 pedestrians from Athens		Among 52 pedestrians from Munich	
Driver Information	Yes (pedestrians)	Yes (%)	Yes (pedestrians)	Yes (%)	Yes (pedestrians)	Yes (%)
Watching	21	31.34	17	26.98	8	15.38
Eye Contact	14	20.90	7	11.11	16	30.77
Hand Gesture	15	22.39	0	0.00	6	11.54
Nodding	15	22.39	0	0.00	4	7.69
Head Movement to The Side	10	14.93	3	4.76	1	1.92
None	31	46.27	35	55.56	27	51.92

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

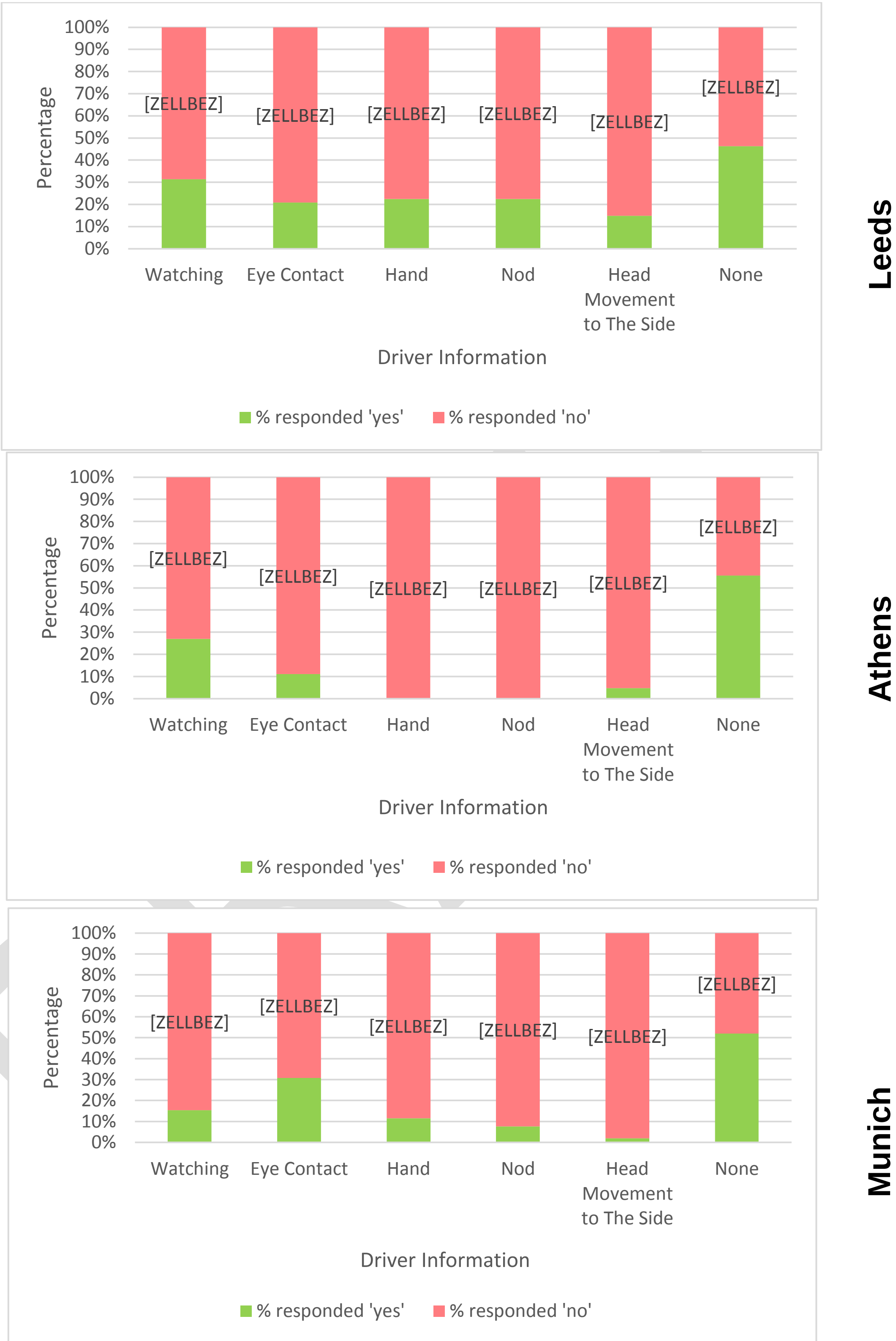


Figure 32: Percentage (%) of pedestrians reported 'yes' (green) and 'no' (red) for each of the driver information used in Leeds (top), Athens (middle) and Munich (bottom). Numbers written on the red bars indicating the number of pedestrians who responded 'no' in each.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

In Leeds, 36 out of 67 pedestrians reported that they use driver information (53.73%) and data shows that the mostly used driver information is 'watching' the driver. In Athens, 28 out of 63 pedestrians reported that they use driver information (44.44%) and data shows that the mostly used driver information is also 'watching' the driver. In Munich, 25 out of 52 pedestrians reported that they use driver information (48%) with the mostly used information as 'eye contact'.

Average of driver information used by each pedestrian were calculated by taking the sum of scores for each driver information used and divided by 5 (watching, eye contact, hand, nod, and head movement to the side). One-sample t-test revealed that the average of driver information used was significantly higher than 0 for Leeds, $t(66) = 7.28$, $p < .001$; comparable results were found in Athens, $t(63) = 6.42$, $p < .001$ and Munich, $t(51) = 5.82$, $p < .001$.

In Leeds, a paired-sample t-test revealed that pedestrians were significantly higher in reporting using vehicle information (M=34.33%, SD=15.94) to decide whether it was safe to cross than using driver information (M=18.81%, SD=21.14), $t(66) = 6.15$, $p < .001$. Same was found for Athens, $t(63) = 3.68$, $p < .001$: vehicle information (M=14.51%, SD=8.70) and driver information (M=8.57%, SD=10.60) and in Munich, $t(51) = 6.29$, $p < .001$: vehicle information (M=35.77%, SD=23.54) and driver information (M=13.46%, SD=16.67).

Indicating Intention – Pedestrians indicated what kind of intention information was provided to show their intention of crossing from each locations (see Table 9) and they were allowed to choose more than one from the options. Figure 33 provides visual representation of each information used by pedestrians in Leeds, Athens and Munich respectively to indicate their crossing intention.

Table 9: Number and percentages of pedestrians who responded 'yes' for each intention information provided from each location

	Among 67 pedestrians from Leeds		Among 63 pedestrians from Athens		Among 52 pedestrians from Munich	
Pedestrian Information	Yes (pedestrians)	Yes (%)	Yes (pedestrians)	Yes (%)	Yes (pedestrians)	Yes (%)
Step Forward	11	16.42	29	46.03	23	36.51
Eye	7	10.45	27	42.86	9	14.29
Hand	7	10.45	0	0	5	7.94
Head Movement	21	31.34	21	33.33	10	15.87
Other	1	1.49	1	1.59	15	23.81

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

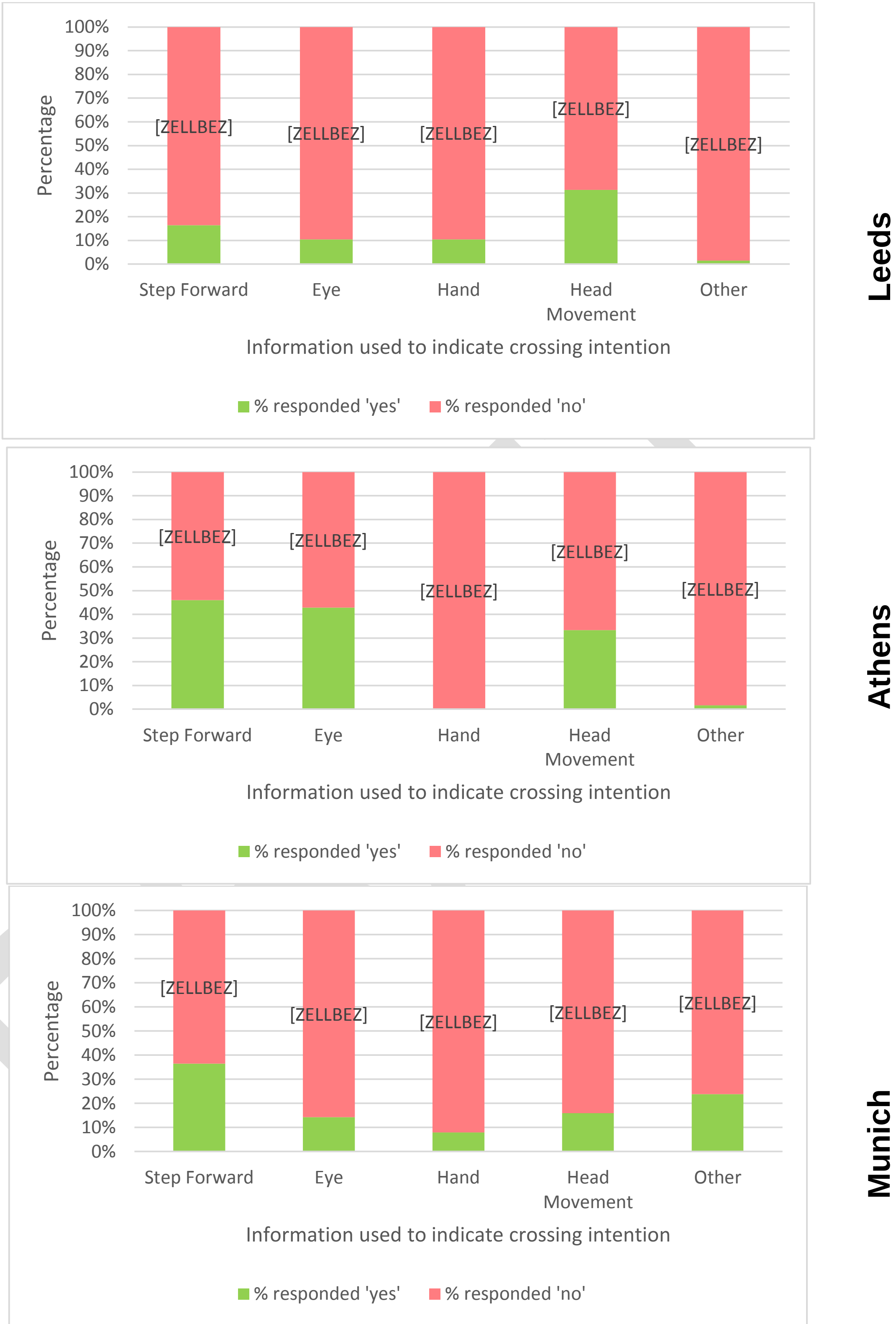


Figure 33: Percentage (%) of pedestrians reported 'yes' (green) and 'no' (red) for each of the information used in Leeds (top), Athens (middle) and Munich (bottom) to indicate their crossing intention. Numbers written on the red bars indicating the number of pedestrians who responded 'no' in each.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Average of information used to indicate crossing intention by each pedestrian were calculated by taking the sum of scores for each information used and divided by 5 (step forward, eye, hand, head movement, other). One-sample t-test revealed that the average of intention information (M=14.03%, SD=14.78) provided was significantly higher than 0 in Leeds, $t(66) = 7.77$, $p < .001$; same results was found in Athens, $t(62) = 18.54$, $p < .001$ and Munich, $t(51) = 17.67$, $p < .001$.

RUBQ – The Adolescent Road User Behaviour Questionnaire was investigated by providing the proportion of pedestrians who responded 'Never', 'Rarely', 'Sometimes', 'Often' and 'Very Often' for each of the questions (see Figure 34 for Leeds, Figure 35 for Athens and Figure 36 for Munich). An average RUBQ score was also calculated for each of the locations. The RUBQ Questionnaire can be found in the Annex.

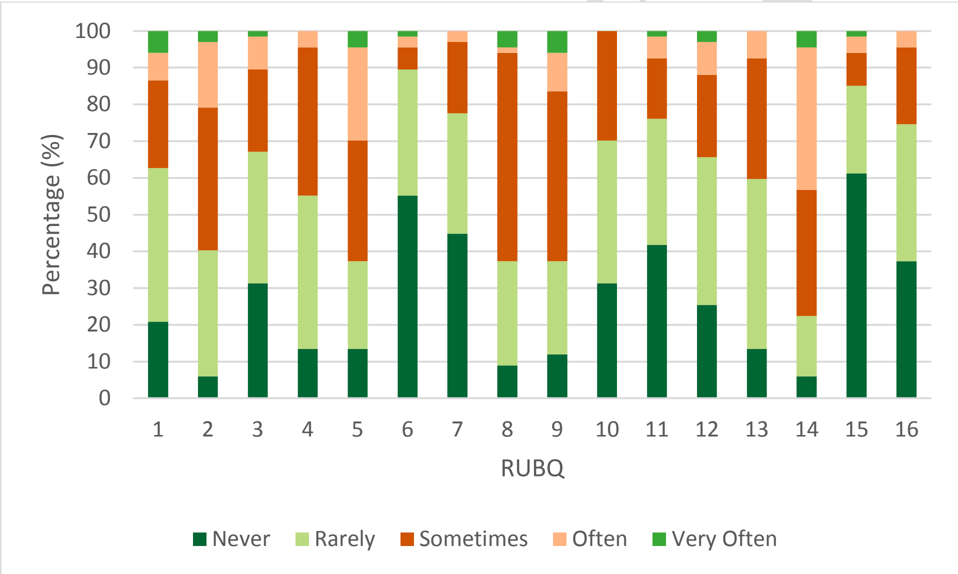


Figure 34: Percentage (%) of pedestrians reported 'Never', 'Rarely', 'Sometimes', 'Often' and 'Very Often' for each of the questions in Leeds

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

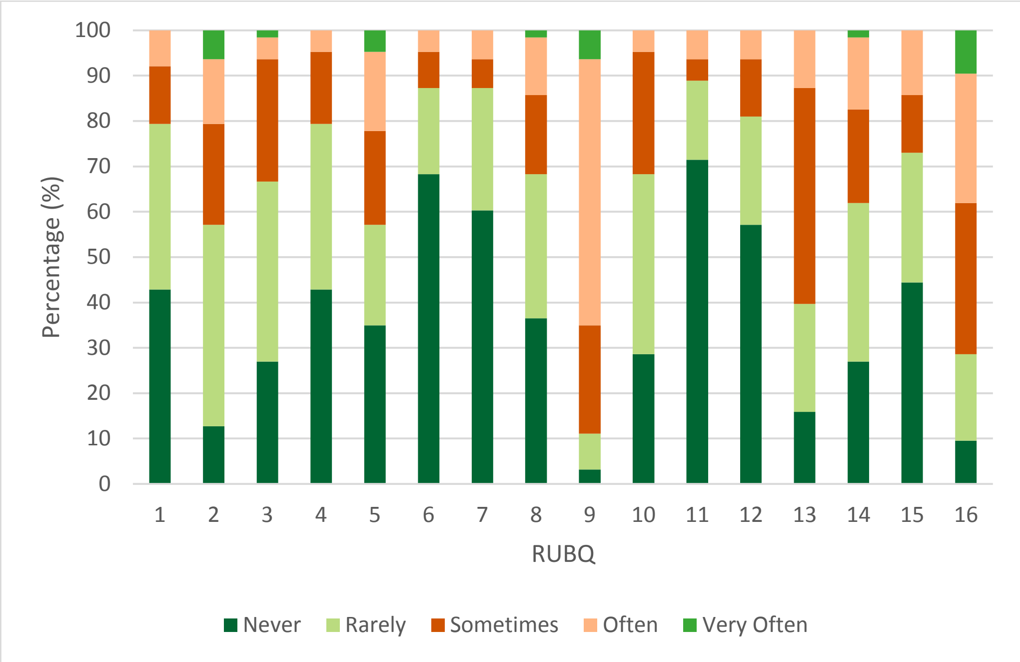


Figure 35: Percentage (%) of pedestrians reported 'Never', 'Rarely', 'Sometimes', 'Often' and 'Very Often' for each of the questions in Athens

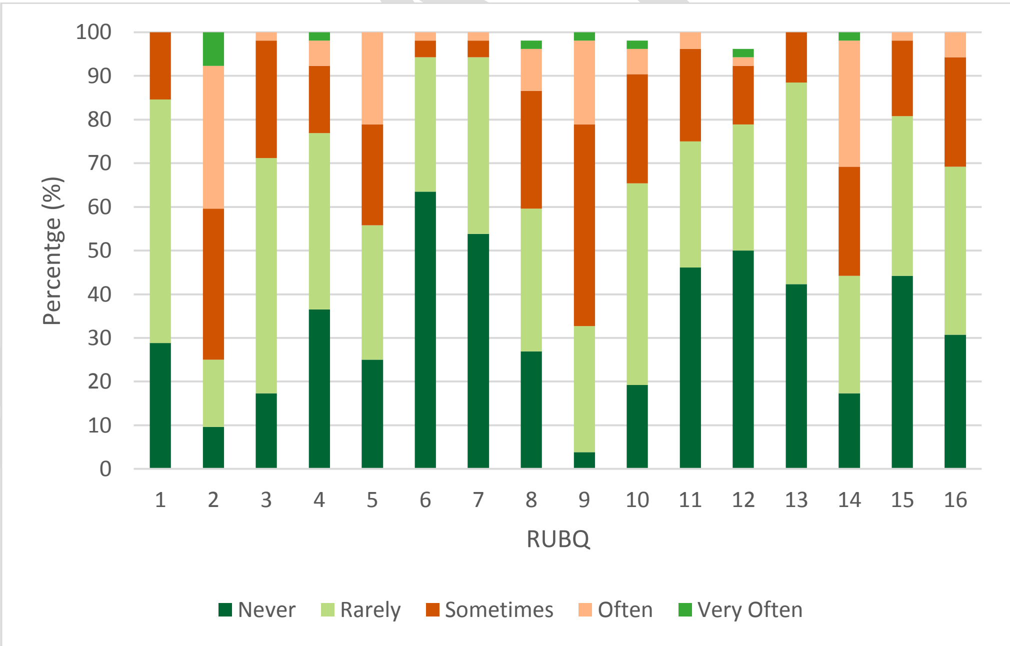


Figure 36: Percentage (%) of pedestrians reported 'Never', 'Rarely', 'Sometimes', 'Often' and 'Very Often' for each of the questions in Munich

The answers from the questionnaires seem to show the same tendencies. Nonetheless an in-depth statistical analysis combining the questionnaires with the protocol data and correlating individual items also evaluating cross-cultural differences will be conducted within Task 2.2. and presented in D2.2.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

6.3 Running Commentary Method

This section details results obtained with the running commentary study, conducted by the ICCS in Athens. While the methodological approach differs from the observational studies, the obtained results are still comparable and yield valuable insights into the perception and decision making of drivers in vehicle-vehicle and vehicle-pedestrian encounters.

6.3.1 Driver-driver interactions relevant to left and right turns

The observed interactions and the observed responses by other drivers per scenario are shown in Table 10. In 146 of the 188 observed left turns and in 126 of the 179 observed right turns, an interaction was started by the subjects. In 62 and 60 cases respectively, a response by the other driver was annotated. The type of the other driver's vehicle is shown in Table 10. In 23 of 25 interactions with drivers of large vehicles, the other driver reacted to the interaction started by the subject. Only 7 out of 58 motorcycle riders reacted to the interaction started by the subject.

Table 10: Interaction starts and other drivers' responses per scenario

	Number of turns	Number of interactions (started by the subjects)	Number of interactions where the other driver reacted
Left turn from 2-way street	188	146 (64 passenger car, 36 taxi, 16 large vehicle, 30 motorcycle)	62 (26 passenger car, 18 taxi, 14 large vehicle, 4 motorcycle)
Right turn to 2-way street	179	126 (63 passenger car, 26 taxi, 9 large vehicle, 28 motorcycle)	60 (33 passenger car, 15 taxi, 9 large vehicle, 3 motorcycle)

The signals or cues by the subjects are shown in Table 11. The row "Nothing" refers to interactions where the subject started turning knowing that the other driver would have to react and slow down. A relevant accompanying comment was "I am sure that he/she has seen me, so I can turn, because I know that he/she can and will yield".

Subjects' edging, use of headlights and gesture/nodding was followed by a response by the other driver in most of the interactions when they were used by the subjects. The turn indicator alone was not so effective, especially for right turns when the other driver, coming from the left of the subject's vehicle, could not perceive the right turn indicator. One subject specifically mentioned his/her intense gazing towards the other drivers as a means to enforce his/her priority on them.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Table 11: Signals or cues by the subject

	Left turn from 2-way street		**Right turn to 2-way street**	
Observed signal / cue by the subject	**Number of started interactions (N=146)**	**Number of interactions with other driver's reaction (N=62)**	**Number of started interactions (N=126)**	**Number of interactions with other driver's reaction (N=60)**
Turn indicator	119	40	66	21
Turn indicator + Edging	17	17	10	10
Turn indicator + Edging + Headlights	2	2		
Turn indicator + Gesture/Nodding	1	1		
Turn indicator + Gesture/Nodding + Edging	1	1		
Edging	1		18	12
Gesture/Nodding			3	2
Nothing	5	1	29	15

The signals or cues by the other drivers are shown in Table 12. The other driver's deceleration or stopping was always followed by the subject turning in front of the other vehicle. The same holds true when the other driver made a gesture/nodded or when the other driver used the turn indicator. The latter because it indicated a change in the other vehicle's trajectory, so no more conflict with the subject's vehicle was expected. The headlights by the other driver did not always result in the subject turning in front of the other vehicle, so the interpretation of this signal is rather done complementary to other signals and cues. Acceleration and use of horn by the other driver was not followed by the subject turning, they were rather interpreted as other's intention to not yield.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Table 12: Signals or cues by the other driver

	Left turn from 2-way street		Right turn to 2-way street	
Observed signal / cue by the other driver	**Number of started interactions (N=146)**	**Number of interactions with other driver's reaction (N=62)**	**Number of started interactions (N=126)**	**Number of interactions with other driver's reaction (N=60)**
Gesture/Nodding			1	1
Headlights	7	4	2	1
Horn	1			
Accelerate			2	
Decelerate	22	22	22	22
Decelerate + Gesture	3	3	2	2
Decelerate + Headlights	1	1		
Decelerate + Headlights + Gesture			1	1
Stop	25	25	24	24
Stop + Gesture	4	4	1	1
Stop + Headlights	1	1		
Stop + Horn	1	1		
Turn indicator			4	3
Opportunity due to other event			3	2
Nothing	81	1	64	2

The sequences of annotated signals / cues for 61 interactions relevant to left turns where the other driver reacted are shown in Figure 37. In the left, the signals and cues by the subjects are drawn, in parenthesis are the number of observations of each signal or cue. Similarly, the right boxes represent the signals or cues by the other drivers. The arrows depict the sequence of actions.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

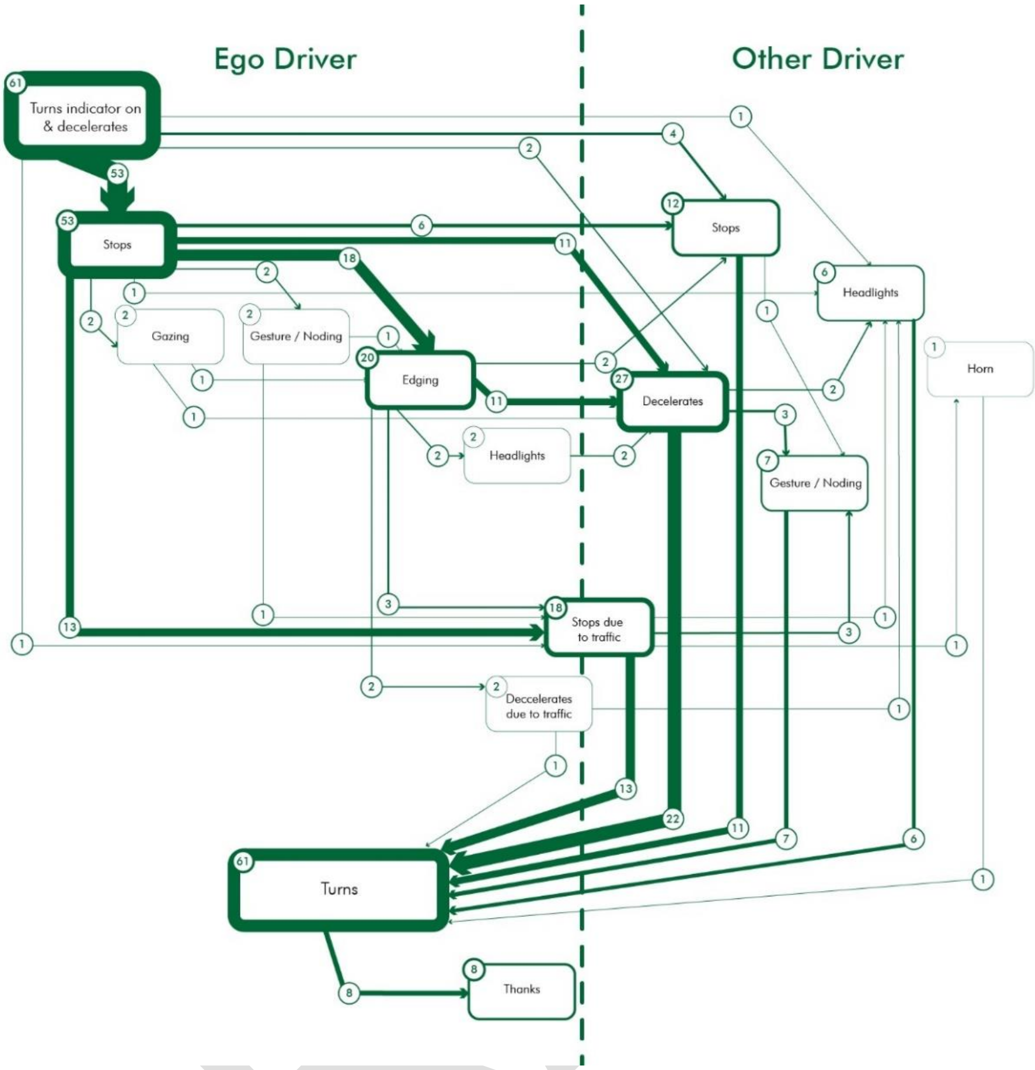


Figure 37: Sequences of observed signals/cues in interactions between drivers relevant to left turns

It seems that a typical sequence of actions for an interaction relevant to a left turn in the specific locations can be described as follows:

- The subject turns the indicator on and decelerates.
- If the oncoming driver reacts and stops the vehicle while the subject decelerates, the subject turns.
- Else, the subject comes to a full stop and waits.
- In these circumstances, the subject frequently edges, namely moves the vehicle a bit forward, possibly trying to make an oncoming driver to yield. Sometimes, the subject flashes headlights to the oncoming driver, makes a gesture, nods or tries to achieve eye contact with the oncoming driver.
- When an oncoming driver decides to yield, he/she decelerates. Sometimes, the other driver flashes headlights or makes a gesture / nods towards the subject.
- Then, the subject turns.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

The sequences of annotated signals / cues for 60 interactions relevant to right turns to two-way street where the other driver reacted are depicted in Figure 38.

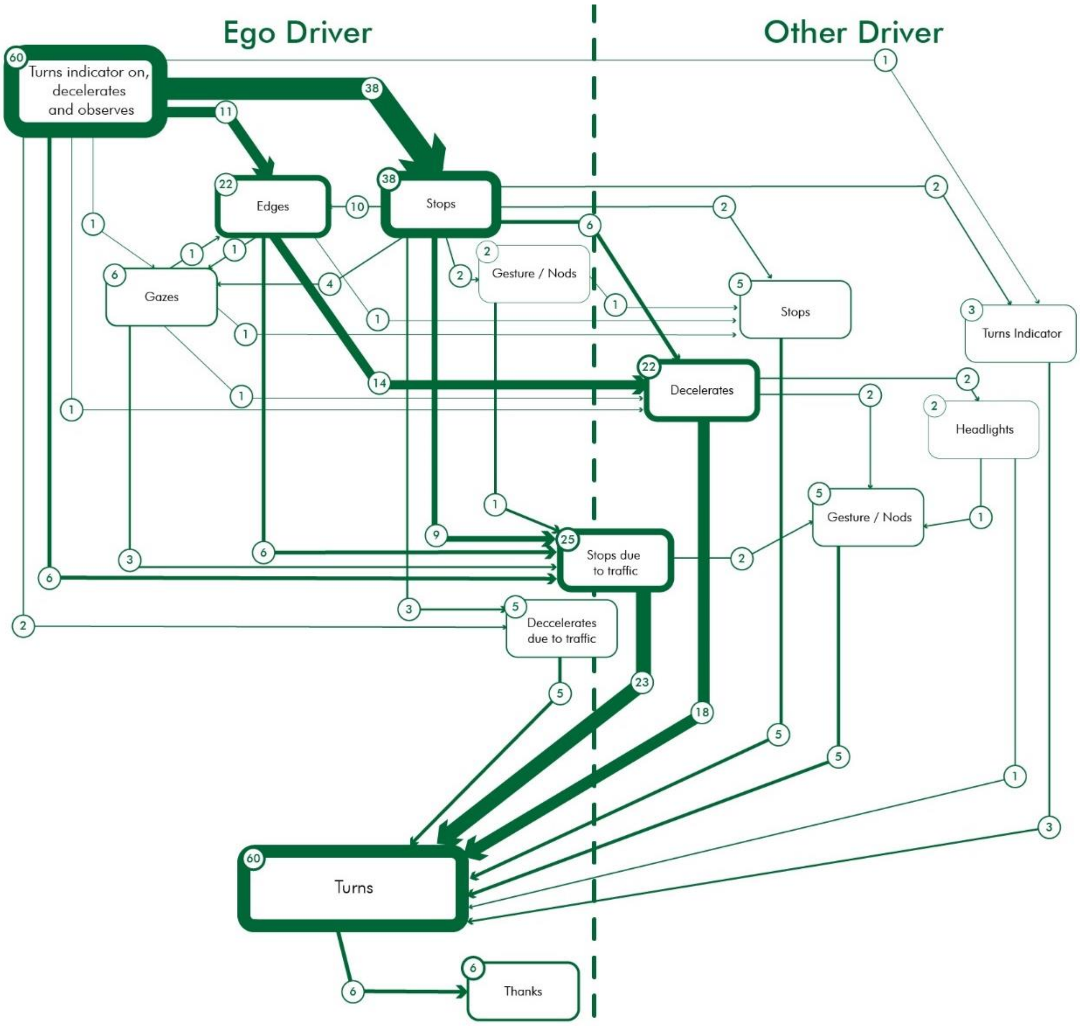


Figure 38: Sequences of observed signals/cues in interactions between drivers relevant to right turns

A typical script for the actions during an interaction relevant to a right turn in the specific location is the following:

- The subject decelerates, turns on the indicator and searches the environment for oncoming traffic. Frequently the subject comes to a full stop.
- Sometimes the subject edges the vehicle a bit forward.
- More rarely, the subject gazes towards the other driver, trying to get in eye contact, and makes gestures or nods.
- When the other driver decides to respond, he/she normally decelerates or stops the vehicle.
- Headlights or gestures/nods from the other driver are rarer.
- Then the subject turns.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

According to commentaries relevant to left turns, subjects consider that achieving eye contact is a good means to convince the other driver to yield. Even the use of headlights is a means to attract the other driver's attention. Additionally, drivers seem to monitor the other drivers' gaze orientation and they plan their behaviour according to whether they believe that the other driver has or has not perceived them. Other cues are used by drivers to anticipate the evolution of the situation. For example, a motorcyclist's foot moving to the ground is interpreted as intention to stop, the presence of people at the bus stop creates the expectation that the bus will stop there.

According to commentaries relevant to right turns, edging is intentionally used by subjects so that the drivers coming from the left can see them. They stated that they have to give a sign to the other driver, without annoying, in order for the other driver to yield. Intense gazing towards the other driver was again considered as a good means to make the other driver yield. On the contrary, when the other driver does not gaze towards the subject although he/she normally should, this intentional avoidance of eye contact is interpreted by the subjects as "he/she will not yield priority". Subjects check whether the other drivers have perceived them or if they are distracted, for example speaking on the phone, and plan their future motion accordingly. For example, if a driver is speaking on the phone, the subjects were more conservative in estimating safe gap to start turning. Additionally, subjects seem to estimate the time that they will need to wait before turning. If they expect that they will not wait long, for example there is only one vehicle from the left and then the street is free, then they wait. If they expect that they will wait long, for example there is a cue of vehicles, they accept shorter gaps to start their turning. Finally, drivers seem to take advantage of opportunities due to external events, for example a pedestrian crossing the street is a green light for the subjects to turn, since they expect that the driver from the left will slow down.

6.3.2 Patterns of driver – pedestrian interaction

In total 487 driver-pedestrian interaction cases were recorded, annotated and analysed from the driver's point of view. These included 265 interactions on straight segments, 80 on left turns from 2way street, 99 on right turns to 2way street and 43 interactions on right turns to 1way street from 1way street. 316 were crossing interactions, 82 were parallel interactions with pedestrian walking in the same direction as the participant's vehicle and 89 parallel interactions with the pedestrian walking opposite to the vehicle. The pedestrian categories are shown in Table 13. The categories are non-exclusive, meaning that there were elderly with bags or standard pedestrians with pets etc. The split in interactions with individual pedestrian and pedestrians in groups is shown in Table 14.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Table 13: Types of pedestrians in Driver-Pedestrian interaction cases (487 cases)

	Crossing	Parallel
Standard	273	154
Elderly	43	19
With shopping Bag	111	56
With Stroller	5	6
Pet	2	3
Children	8	4

Table 14: Interactions with individuals or groups in Driver-Pedestrian interaction cases (487 cases)

	Crossing	Parallel
Individual	255	152
Group	61	19

The above categories are exclusive, meaning that the sum of each column sums up to the total incidents analysed.

A first observation was that a driver-pedestrian interaction can be resolved either through a form of physical movement co-ordination –mediated only by implicit signals e.g. emitted from the pedestrian body movement/head orientation, or it may involve a form of non-verbal communication –mediated by explicit signals emitted from the two road-users e.g. eye-contact, hand gesture and nodding (turning-lights were not considered).

In addition, based on the drivers' video-assisted retrospective commentaries it was evident that, from the driver's point-of-view, interaction cases fell in two broad categories depending on the driver's attention towards the pedestrian or alternatively depending on his/her confidence about the future intended action of a pedestrian. In the majority of interaction cases, a driver would reside on implicit signals (e.g. pedestrian's body movement/head orientation, pedestrian gaze) that were clearly meaningful based solely on situational factors and norms. Instead, in roughly ¼ of the cases drivers exhibited a certain level of uncertainty about the future intended action of a pedestrian. In these cases the driver tended to search for further cues or signals from the pedestrian in order to infer his/her intention and/or to emit a signal to the pedestrian (e.g. flashing lights; hand gesture).

This observation was substantiated by an analysis of drivers' eye-fixations which revealed that in the majority of cases, during interaction, drivers fixated three times or less at the pedestrian concerned (Figure 39). Based on these data, three fixations was taken as a reasonable criterion to nominally divide interaction cases from the point-of-view of the driver into two patterns, routine (Pattern A) and non-routine (Pattern B) ones. In the so-called Pattern A, the driver is aware of the pedestrian but does not allocate his/her full attention to the pedestrian. Instead, in the Pattern B, (as the number of driver

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

fixations suggest), the pedestrian constitutes the driver's primary concern, at least at some point in time during the interaction.

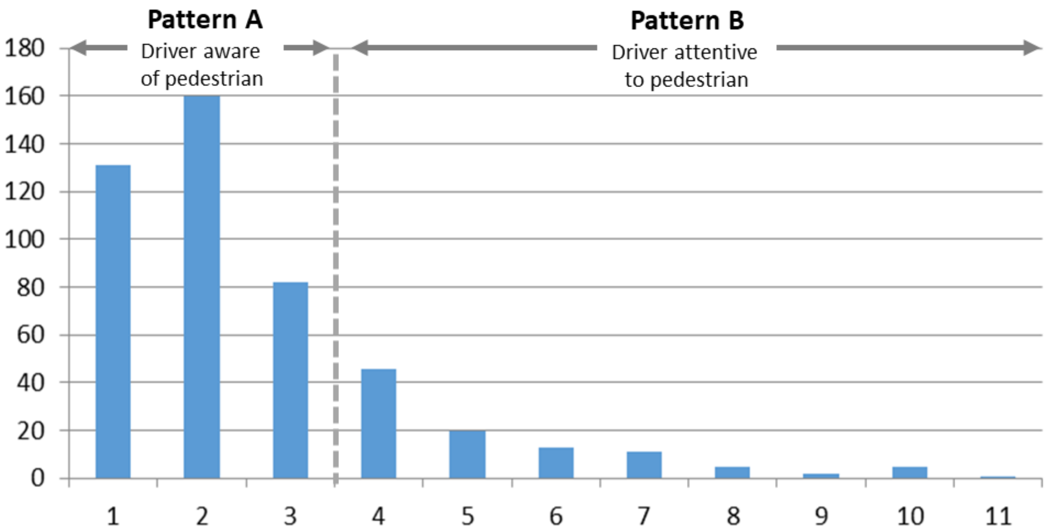


Figure 39: Distribution of interaction cases according to number of driver fixations on pedestrian (1-11). Four fixations or more towards the pedestrian was taken as a nominal criterion signifying that the driver is attentive to pedestrian and/or uncertain of his intent

As seen in Table 15A, the vast majority (77%) of driver-pedestrian interaction cases fall in pattern A, which is characterized by the driver's confidence about the future intended action of a pedestrian based on implicit signals (e.g. pedestrian's body movement/head orientation including gaze) that become meaningful on the grounds of situational factors and norms. Instead, about one fourth of cases (23%) are characterized as Pattern B, suggesting a level of driver's uncertainty about the future intended action of a pedestrian that might lead the driver to emit a signal to the pedestrian (e.g. flashing lights; hand gesture) and/or to search for further signals from the pedestrian in order to infer his/her intention.

Nevertheless, in both patterns, the main issue of concern for the driver is to clarify the pedestrian's intended future action, which in turn affects the driver's decision to give priority to the pedestrian or not. Therefore, all cases were labelled according to the interaction outcome as (i) Pedestrian passing first, (ii) Driver passing first and (iii) Non Applicable (N/A signifying cases of parallel movement with mutual yielding).

Table 15: Relative frequencies of passing road user (D: ego=driver; P: pedestrian; N/A: non-applicable) when a driver is confident about the future intended action of a pedestrian (Pattern A) or less confident (Pattern B), and type of signals emitted by the pedestrian and driver. 1A: driver-pedestrian interaction in all road sections (N=487); 1B: driver-pedestrian interaction at crossings (N=316); 1C: driver-pedestrian interaction parallel to the road (N=171)

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Table 16 (A, B and C): Driver-Pedestrian interactions

A: Driver-Pedestrian interaction totals (487 cases)								
Signal(s) from Pedestrian	**Pattern A** **Driver aware of pedestrian**				**Pattern B** **Driver focuses on pedestrian**			
	N (376)	Passing road user				Passing road user		
		D	P	N/A	N (111)	D	P	N/A
Implicit signals	**376**	**0.29**	**0.44**	**0.27**	***46***	**0.24**	**0.54**	**0.22**
[1] = Body movement/ orientation	219	0.21	0.45	0.35	25	0.08	0.60	0.32
[1]; Gaze	157	0.41	0.42	0.17	21	0.43	0.48	0.10
Explicit signals					**65**	**0.20**	**0.71**	**0.09**
[1]; Eye-contact					52	0.17	0.75	0.08
[1]; Eye contact; Hand gesture/ Nodding					13	0.31	0.54	0.15
Hand gesture / Nodding from Driver					24	0.13	0.83	0.04

B: Driver-Pedestrian interaction while pedestrian crosses (316 cases)								
Signal(s) from Pedestrian	**Pattern A** **Driver aware of pedestrian**				**Pattern B** **Driver focuses on pedestrian**			
	N (242)	Passing road user				Passing road user		
		D	P	N/A	N (74)	D	P	N/A
Implicit signals	**242**	**0.35**	**0.65**		**23**	**0.13**	**0.87**	**-**
[1] = Body movement/ orientation	124	0.26	0.74	-	12	0.00	1.00	-
[1]; Gaze	118	0.45	0.55	-	11	0.27	0.73	-
Explicit signals					**51**	**0.18**	**0.82**	**-**
[1]; Eye-contact					42	0.14	0.86	-
[1]; Eye contact; Hand gesture/ Nodding					9	0.33	0.67	-
Hand gesture / Nodding from Driver					21	0.14	0.86	-

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

C: Driver-Pedestrian interaction while pedestrian moves parallel to the road (171 cases)								
Signal(s) from Pedestrian	**Pattern A** **Driver aware of pedestrian**				**Pattern B** **Driver focuses on pedestrian**			
	N (134)	Passing road user				Passing road user		
		D	P	N/A	N (37)	D	P	N/A
Implicit signals	***134***	**0.19**	**0.05**	**0.76**	**23**	**0.35**	**0.22**	**0.43**
[1] = Body movement/ orientation	95	0.14	0.06	0.80	13	0.15	0.23	0.62
[1]; Gaze	39	0.31	0.03	0.67	10	0.60	0.20	0.20
Explicit signals					**14**	**0.29**	**0.29**	**0.43**
[1]; Eye-contact					10	0.30	0.30	0.40
[1]; Eye contact; Hand gesture/ Nodding					4	0.25	0.25	0.50
Hand gesture / Nodding from Driver					3	0	0.66	0.33

Findings related to the all interaction cases

- 87% of all interaction cases (N=422) were resolved solely through implicit signals (i.e. through body movement, head/body orientation and gaze).
- No gaze from pedestrian (N=230) tends to result in Pedestrian passing first (P=0.45; D=0.20; N/A=0.35). This is particularly evident in Pattern B crossing cases with no gaze (N=12) where the driver yields in 100% of cases.
- On the other hand, (perceived) gaze from pedestrian but without eye-contact (N=178) tends to promote Driver passing first (P=0.43; D=0.41; N/A=0.16).
- Findings related to Pattern B interaction cases (i.e. with > 4 driver's fixations on the pedestrian)
- Irrespective of pedestrian's gaze, Pattern B cases (N=111) tend to result in pedestrian passing first (P=0.64; D=0.21; N/A=0.14).
- 88% of eye-contacts between pedestrian and driver (N= 65) occur when the Pedestrian is on street surface (not on pavement).
- From the 111 Pattern B cases, 65 (59%) resulted in eye-contact between pedestrian and driver and a further 28 cases (24%) resulted in signalling through gestures (24 driver gestures, 13 pedestrian gestures and 9 mutual).

An illustration of the outcome of these interactions are presented in Figure 40.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

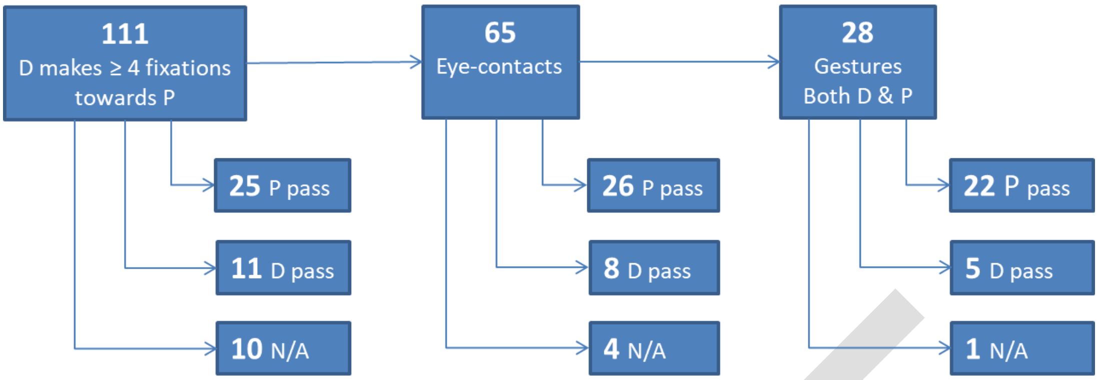


Figure 40: Pattern B V-P interaction cases (i.e. with ≥4 driver eye-fixations on pedestrians) and related outcomes

As can be derived from Figure 40, eye-contact between pedestrian and driver (N= 65) enhances the effect of pedestrian passing first (P=0.71; D=0.20; N/A=0.09). In addition, a gesture/nodding between pedestrian and driver (N=28) enhances even more the effect of pedestrian passing first (P=0.79; D=0.18; N/A=0.03).

Findings related to crossing cases (N=316)

- In crossing cases with no gaze from pedestrian (N=124) there is a marked tendency of pedestrian passing first (P=0.74; D=0.26)
- Multiple (>4) driver's fixations on pedestrian (N=74) resulted in a marked tendency of pedestrian passing first (P=0.84, D=0.16)
- No gaze from pedestrian plus multiple (>4) driver fixations on pedestrian (N=12) resulted in 100% pedestrian passing first (P=1, D=0).
- When a pedestrian initiated movement while approaching (N=40), this resulted in 100% pedestrian passing first (P=1, D=0).
- When a pedestrian kept pace (N=140), this resulted in a marked tendency of Pedestrian passing first (P=0.93, D=0.07).
- When the pedestrian either stopped, slowed down or stepped back (N=68), this resulted in a tendency of Driver passing first (P=0.26; D=0.72; N/A=0.02).
- When the pedestrian remained idle at the start of the interaction (N=57), this resulted in a tendency of Driver passing first (P=0.32; D=0.63; N/A=0.05)
- When a pedestrian speeded-up at the start of the interaction (N=11), this resulted in 100% pedestrian passing first (P=1, D=0)

Findings related to interaction sequencing

Sequence diagrams for left and right turns were produced designating implicit and explicit signals emitted between pedestrian and drive.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

All sequences were mapped in a “four phase communication grid” as in Table 17 below.

Table 17: Interaction phases in the sequence diagrams

Phase	Road user	Description
Phase I.	Pedestrian	Pedestrian movement and cues, followed by gaze at the start of the interaction
Phase II.	Driver	Driver’s behaviour at the start of the interaction along with any signals emitted from the driver
Phase III.	Pedestrian	Any change in pedestrian’s behaviour during the interaction along with any signals emitted from the pedestrian
Phase IV.	Driver	Any final change in the driver’s behaviour or signal emitted

All the interactions started after a certain signal from the pedestrian either implicit or explicit. There are implicit signals (body movement, head body orientation, gaze) and explicit signals (eye contacts, gesture/nodding). The communication begins when the driver responds with either an implicit or explicit signal. Many communications end at this point, while others continue with a response from the pedestrian and a second response from the driver.

The red arrows show the last interaction point (signals) along with the resulting outcomes. There are 3 options: 1) Pedestrian passed in front of the car, 2) Driver passed in front of the pedestrian and N/A meaning that whoever passed in front of the other did not affect the movement of the other (due to long distance or change of direction).

The numbers on the signals’ box or the arrows depicts the number of the reported signals. Also the line thickness is proportional to this number.

This report is part of interACT project that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 723395.

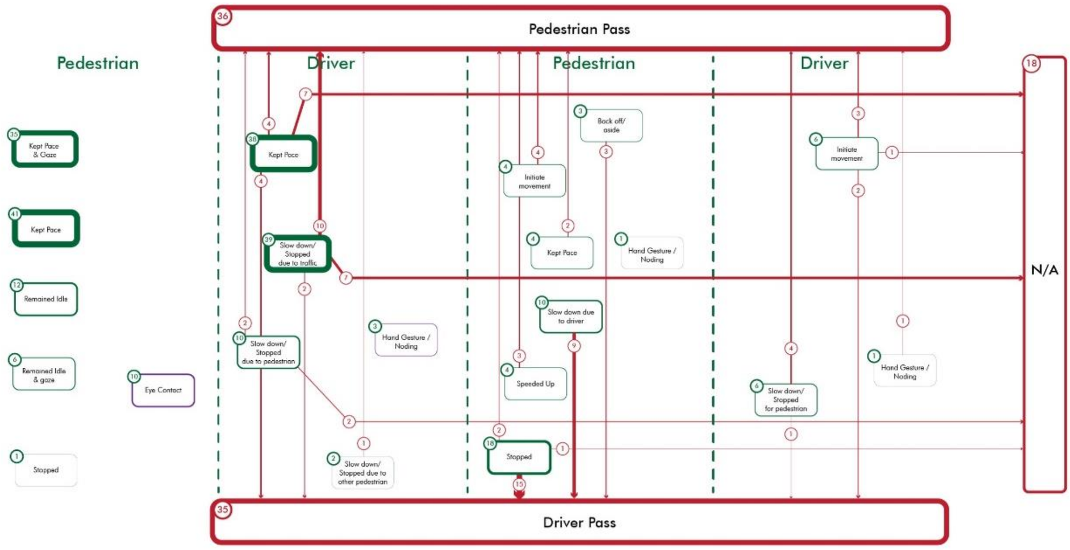


Figure 41: Sequences of observed signals/cues in interactions between drivers and pedestrians in driver's left turns

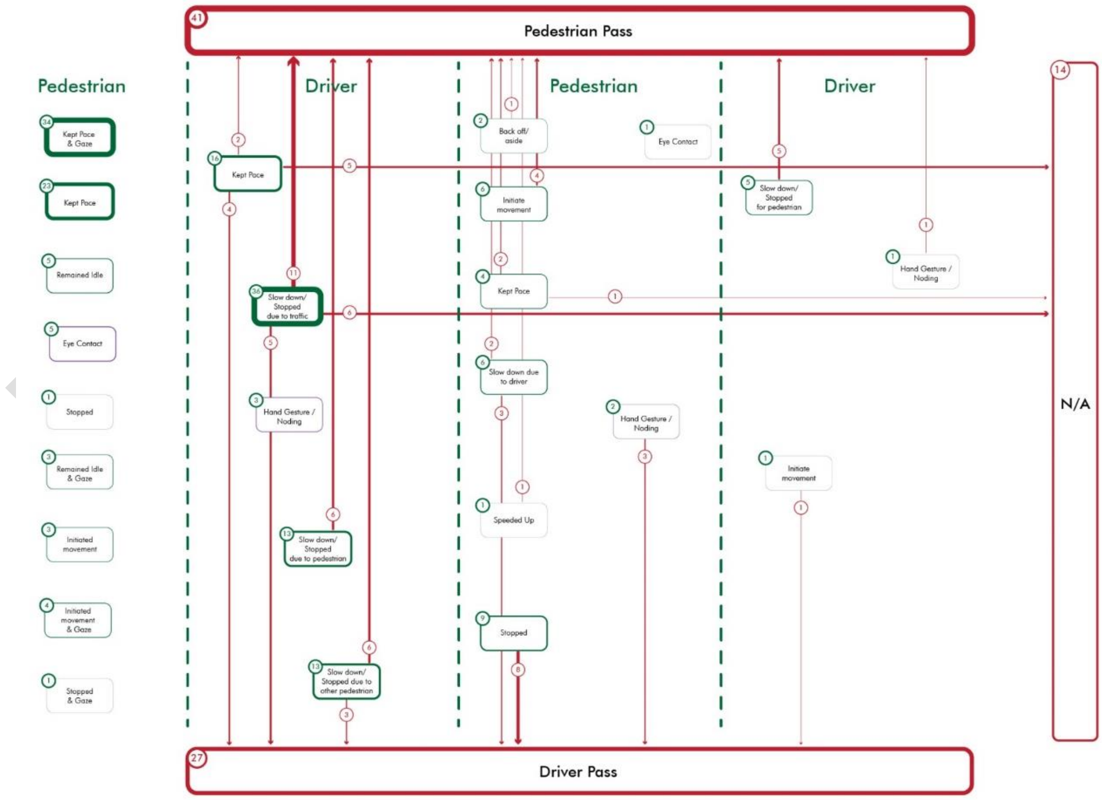


Figure 42: Sequences of observed signals/cues in interactions between drivers and pedestrians in driver's right turns

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

- In 122 cases where the driver kept pace during the interaction sequence, only 12 cases resulted in the pedestrian passing in front (P=0.10; D=0.57, N/A=0.34).
- In 80 cases where both Driver and Pedestrian kept pace at the start of the interaction, only in 9 cases the pedestrian passed in front (P=0.11; D=0.50, N/A=0.39.
- In 198 cases where the Driver slowed down or stopped (both due to traffic, or pedestrian), there is a marked tendency of Pedestrian passing first (P=0.71; D=0.21; N/A=0.08).
- In both left & right turns the “eye-contact” causes a very frequent stopping of the driver for the pedestrian.
- In right turns the driver stopped much more frequently for the pedestrians (the one interacting with and others).
- An interaction during a driver’s left turn began more frequently with pedestrians keeping their pace.
- All recorded hand gestures and nodding were emitted to give priority to the other and never to claim priority.
- When the communication lasted for long usually the pedestrian passed first.
- Most communications end was based only on implicit signals.

It is estimated that the above indices can help predict interaction behaviour by roughly 2/3. Other factors that influence behaviour that were not considered in the present analysis are (i) road user and (ii) road environment particularities, both intrinsic and situational. Table 18 provides examples of these factors.

Table 18: Road User and Road environment factors affecting behaviour that were not considered in the present analysis.

Factor	Road user	Road environment
Intrinsic factors	Risk proneness, agility, perception	Physical barriers, pavement
Situational factors	In a hurry, carrying load, absent minded	Obstructions, other road users, night

This report is part of interACT project that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 723395.

6.4 Models – Quantitative approach

6.4.1 Background and modelling overview

One of the goals of the interACT WP2 is to develop quantitative models of human road user behaviour. These models should capture how individuals interact with other road users, including automated vehicles (AVs). A key application of such models is *virtual testing*, in which computer simulations of AVs and other road users are used to optimise the AV's behaviour, as well as test vehicle safety. While human road user models are currently in use in vehicle testing (Chen, Zhao and Peng, 2017; Kesting, Treiber and Scho, 2005), they generally consider behaviour at the level of traffic microsimulation (i.e. simulate trajectories of road users from equations of motion, while taking into account other road users trajectories). In order to operate safely in mixed traffic environments, AVs will need to interact in complex and robust ways with other road users, at a more detailed level than what is described by current microsimulation-level models. Critically, these interactions will depend on the fine-grained details of human perception, action and decision making

Several decades of research in Psychology and Neuroscience has examined how humans make decisions using the noisy sensory information available to them. A plethora of studies, using laboratory based perceptual tasks, suggests that humans make decisions by accumulating noisy sensory information to a threshold (Ratcliff, 2016). The evidence accumulation hypothesis suggests that people make a decision by accumulating evidence for alterative choices over time, making a decision when the accumulated evidence sufficiently favours one of the alternatives (Bogacz et al., 2006). This accumulation process is subject to random fluctuations, leading to a certain level of stochasticity in the decision making process. This decision process can be described by accumulator models (also known as drift-diffusion models) (Ratcliff, 2016). These have been shown to successfully capture reaction time and accuracy data across a broad range of sensorimotor tasks, including responses to stimuli in traffic, such as brake lights and visual looming of approaching vehicles (Markkula et al., 2016; Svard et al., 2017).

While accumulator models have been invaluable in understanding the behavioural and neurophysiological mechanisms of decision making, they are only applicable to simple low-level decisions (e.g. "I need to brake harder"). However, many of the decisions made by human road users (e.g. is it safe to cross the road?) can be seen as a synergistic ensemble of lower-level decisions. For example, in the case of road crossing, a pedestrian may be required to make several decisions regarding an approaching vehicle (e.g. "do I have time to cross?", "is the vehicle decelerating?", "is the vehicle letting me pass?"), which all feed into the final decision as to whether to cross. Thus in their typical form, accumulator models seem unable to account for the complex high-level decisions involved in human traffic interactions. However, in interACT WP2 (already published by Markkula et al., 2018) we have investigated whether these models could be extended to capture the multiple low-level decision processes inherent in tasks such as road crossing.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

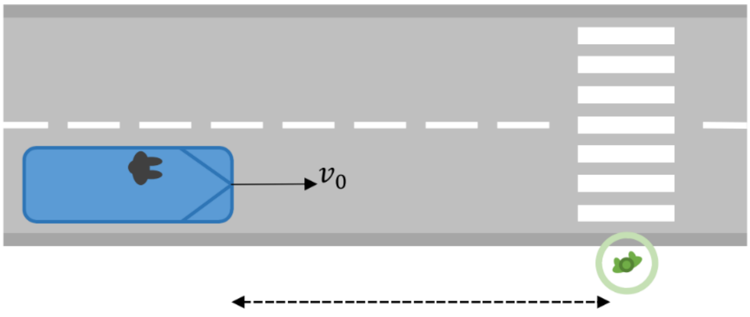


Figure 43: Pedestrian crossing scenario. A pedestrian crosses at an unsignalised junction while a vehicle approaches. Reproduced from Markkula et al., 2018 with permission from SAGE Publications, Inc.[5]

The models described by Markkula et al. (2018), could theoretically be applied to numerous human traffic interactions. For the purpose of this document, we limit of discussion to road crossing scenario, in which a pedestrian decides whether to the cross a road at an unsignalised junction, while a vehicle approaches (see Figure 43). In our modelling framework we view the decision to cross the road as being informed by a number of lower level perceptual decisions (shown in Figure 44). Each of these sub decisions are accumulator models (referred to here as decision modules) which can take multiple sources of sensory information as inputs. In addition, the outputs of each decision module can provide inputs to other modules. The outputs of these lower level modules (the result of a drift diffusion process) then feed into a final "action" module which accumulates these inputs to provide a final decision to cross the road.

As documented by Markkula et al. (2018) with careful parameterisation the model is able to capture qualitative trends reported in the literature. Specifically the distribution of crossing times in relation to a given vehicle approach was found to reflect findings in the literature showing that a proportion of pedestrians wait for an approaching car to come to a complete stop before crossing, while others begin crossing earlier. This occurs in the model because sometimes enough evidence is accumulated to suggest that there is sufficient time to cross the road before the vehicle arrives. However, in some cases the noise present in the model results in the accumulation not reaching threshold before the car is too close to allow safe crossing. In this case the model does not trigger a crossing action until the other decision modules eventually push the action module to threshold (i.e. when the car starts to decelerate or indicates to the pedestrian that they can cross).

[5] http://journals.sagepub.com/home/trr

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

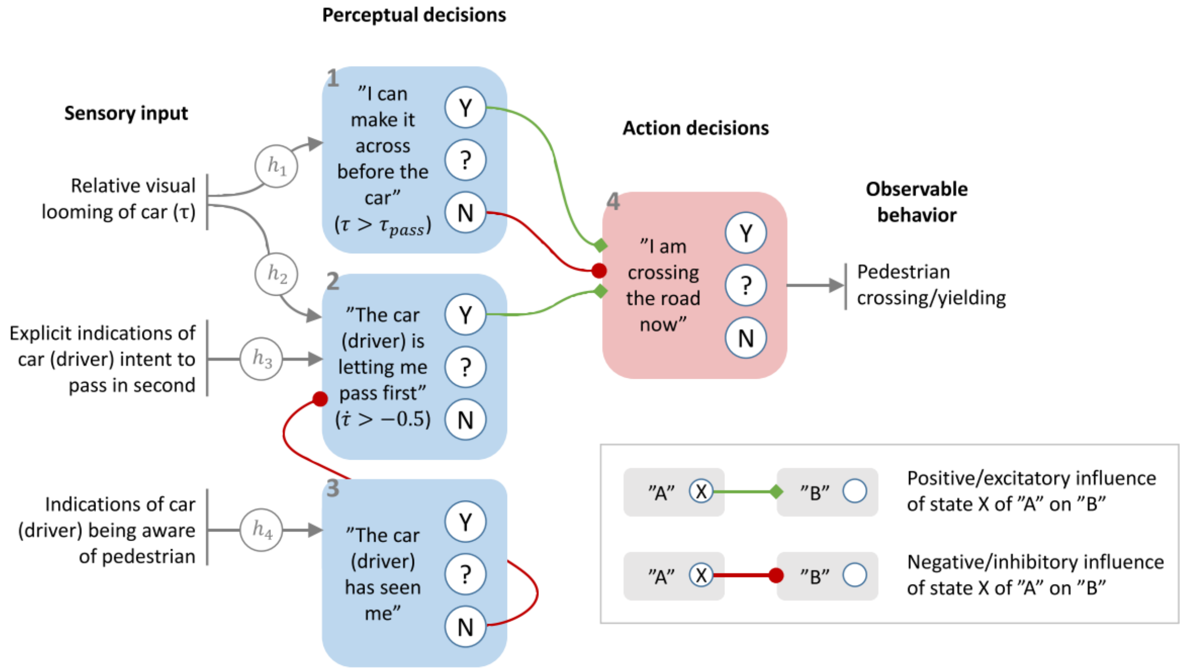


Figure 44: Schematic representation of the pedestrian crossing model. Reproduced from Markkula et al., 2018 with permission from SAGE Publications, inc.[6]

We also explored how the modelling framework could be used to optimise the behaviour of an AV approaching a pedestrian at a crossing. Here we examined how different deceleration profiles could be used to minimise the time lost when stopping at a pedestrian crossing. We found that in general, decelerating faster resulted in the simulated pedestrian starting to cross the road earlier in the AVs approach, which resulted in less time being lost at the junction (see Figure 45). In addition, we found that providing explicit communication from the AV, indicating that it indented to stop, further reduced the time lost. However, this interacted with the effect of changing the deceleration profile, such that decelerating quicker had less effect of the pedestrian's crossing time when the vehicle was providing an explicit communication signal.

[6] http://journals.sagepub.com/home/trr

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

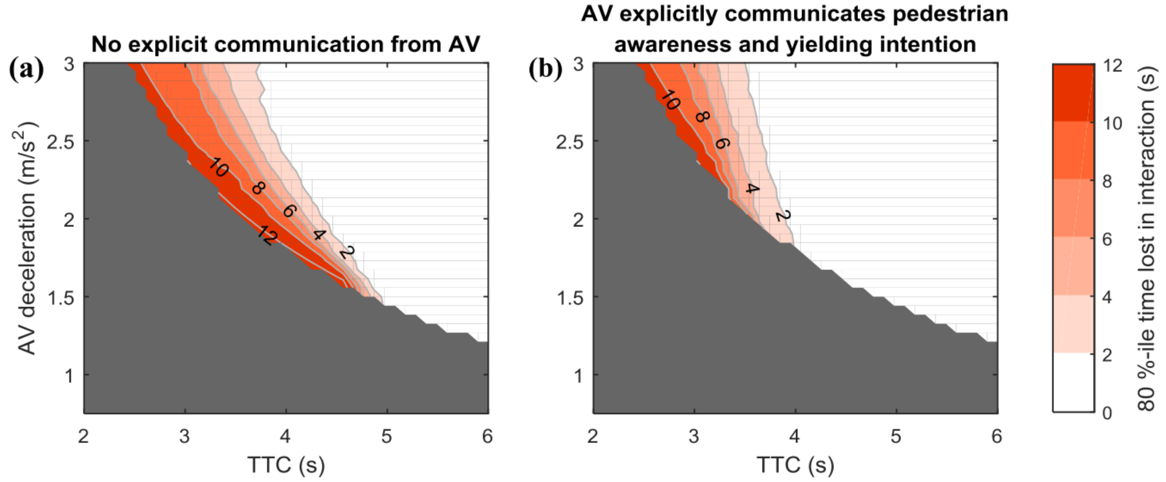


Figure 45: Estimation of time lost at pedestrian crossing as a function of time to contact (TTC) at the time of pedestrian presentation. a) A case with no explicit communication from the vehicle. b) A case with explicit communication. Reproduced from Markkula et al., 2018 with permission from SAGE Publications, Inc.[7]

Model Fitting

The modelling framework outlined above and in (Markkula et al., 2018) appears to capture qualitative aspects of crossing behaviour reported in the literature. The framework also provides a promising approach for optimising AV behaviour through simulations. However, the next challenge is to parameterise the models using quantitative data from real pedestrian crossings. This will allow us to systematically investigate the model's ability to capture human crossing decisions, and provide realistic simulations of crossing behaviour.

The models can be fit using data collected from either real world observations or controlled laboratory experiments. The benefit of laboratory experiments is that we are able to carefully control the environment and the visual information available to the road user. This also allows for the trajectories of the vehicles to be precisely controlled and the geometry of the scene can be carefully defined. Real world observations are more challenging to work with given the complexity and diversity of the pedestrian-vehicle interactions and limitations in measuring the geometry of the scene. The benefit of real world data is that it provides a level of realism and ecological validity which may not be achievable in laboratory experiments. InterACT have already begun providing both kinds of data sources, which are both invaluable for fitting the pedestrian models.

The first challenge in fitting the pedestrian models to either real world or simulator data is finding a suitable method for parameterising the model. Models which capture the mechanistic processes underpinning psychological processes can often be difficult to parameterise because it can easily become difficult or impossible to define a likelihood function (a function which states the probability of a dataset given a particular model parameterisation). The absence of a likelihood function makes it

[7] http://journals.sagepub.com/home/trr

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

difficult to employ optimisation methods such as maximum likelihood estimation or Bayesian estimation approaches. One possible method is to numerically approximate the likelihood function by simulating a large number of samples from the model, and then use an appropriate method to estimate the likelihood of the dataset (e.g. using Kernel Density Estimation). Miletić et al., (2017) showed that this method could be used to successfully fit a leaky accumulator model, using both genetic optimisation algorithms and Markov Chain Monte Carlo (MCMC) methods for Bayesian estimation. However, in practise this method is extremely data inefficient and can take an impractically long time to fit.

Another promising approach in cases where the likelihood function is not known is Approximate Bayesian computation (ABC). While Bayesian estimation techniques, such as MCMC, attempt to sample directly from the posterior distribution, ABC methods are a collection of algorithms which are able to sample from an approximation of the posterior. This is made possible by generating many simulated datasets for possible model parameters (for an overview see Toni et al, 2009). Like Bayesian estimation, ABC can be particularly useful for model fitting as it provides complete distributional information over model parameters. Unlike maximum likelihood estimation, which provides a point estimate of the model parameters which maximise the probability of the data, Bayesian estimation tells you how likely every possible set of model parameters is given the data.

Here we investigated the use of Sequential Monte Carlo ABC (ABC-SMC), an efficient ABC algorithm (Toni et al., 2009), for fitting our pedestrian models. To do so we generated a simulated dataset with 100 crossing observations using the model employed by Markkula et al., (2018). We then attempted to recover the model parameters using the ABC-SMC algorithm. The true parameter values and posterior estimates obtained by the ABC-SMC algorithm are shown in Table 19. The true values always fell within 1 SD of the posterior mean estimate suggesting that we were able to successfully recover the model parameters.

Table 19: Posterior estimates of model parameter values

Parameter	True Value	Posterior Mean	Posterior SD
T	2	3.18	2.62
k_1	0.5	0.59	0.52
k_2	0.5	0.95	1.23
k_3	4.0	2.46	2.16
σ	0.5	0.49	0.17

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Next we performed a posterior predictive check in which we generated new datasets from the approximate posterior distribution. If the ABC-SMC algorithm successfully recovered the real model parameters we should expect the new datasets to look similar to the original data. Figure 46 shows the original data as a histogram, with the replication datasets shown as superimposed kernel density plots. It is clear that our replication data looked a lot like the original data, suggesting that we were able to recover sensible parameters.

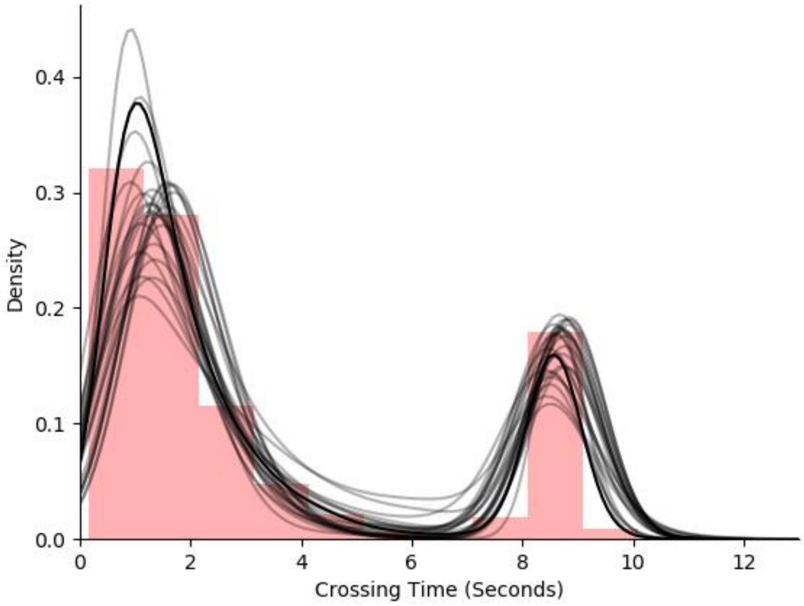


Figure 46: Posterior predictive check. The red histogram shows the real dataset, while the black lines show KDE plots of the datasets simulated from the posterior distribution. The simulations from the posterior match the real dataset well.

Future work

We now have a modelling framework which is able to capture the qualitative aspects of pedestrian crossing behaviour which are observed in the literature. In addition, we have developed a pipeline for fitting these models to data from human subjects by exploiting recent advances in Approximate Bayesian Computation. Our next step as part of the interACT project is to begin fitting the models to both the natural observation data collected in the UK, Germany and Greece. In addition, we are working with partners in Japan as part of the SIP-Adus project (http://en.sip-adus.jp/rd/), who have collected similar datasets on Japanese roads. Thus we will be able to test whether our models are able to account also for any differences in behaviour between Europe and Japan. This may be critical for virtual testing, where simulations may not only need to account for cross country differences in road rules and regulations, but also cultural differences in crossing behaviour.

In addition to the real world data, interACT WP2 and WP4 are in the process of preparing and running various carefully controlled laboratory experiments using virtual reality (VR). This will provide us with

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

the ability to precisely control the visual stimuli available to test subjects, allowing us to explore explicit hypothesis regarding road crossing behaviour. We are currently developing a series of studies at TUM and the University of Leeds which will provide this rich laboratory dataset. These can then be tested and validated using the interACT WP2 naturalistic observation datasets.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

7. Generalizable Findings

In general, urban traffic is highly dynamic: road users try to continuously carry on towards their destination following traffic regulations. If this is not possible due to increased traffic density, congestions or other hindrances, road users will adapt to these situations. Interactions occur if the normal traffic flow is impaired and usually consists of negotiating the right of way between at least two road users. While traffic participants generally try to avoid interaction-demanding situations and conflicts, some situations require communication and cooperation to achieve a certain goal. Within the four observed use cases (see Ch. 3.2) the type of involved traffic participants influenced the way an interaction took place. Hence, the generalizable findings are clustered in Vehicle-Pedestrian and Vehicle-Vehicle encounters combining the locations.

The following chapter aims to give an understanding of the observed use cases by reflecting the subjective views of the observers in the three involved countries. Recommendations for the development of the CCPU are based on human road user behaviour, and thus not universally valid, as automated vehicles will likely have significant influences on traffic, once they are introduced onto urban roads.

7.1 Vehicle-Pedestrian encounters

Interactions, where explicit communication is utilized, occur rarely in pedestrian vehicle encounters. By behavioural adaptation of either involved road user, most potential interaction-demanding situations are resolved before they actually form. This means that the AV’s intelligence (CCPU within interACT) has to identify potential encounters early and try to resolve them by adapting the driving behaviour in a way that the other road user understands the intention of the vehicle without utilizing any explicit communication. Within the observations, explicit communication from drivers towards pedestrians was used, when the kinematic adaptation did not result in the anticipated behaviour and the relative velocity and distance was very low.

In the observed use cases interactions occurred at low velocities and mostly revolved the question of “who goes first?”. In almost all cases, observed interactions could be described as traffic participant one (TP1) either yielding its right of way to TP2 or, in case that TP2 has the right of way, reassuring TP2 that TP2 can go first. In very few cases this communication approach failed, resulting in somewhat ambiguous situations where warning messages were utilized (i.e. honking).

This report is part of interACT project that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 723395.

Example: A pedestrian approaches a road, which he intends to cross, and looks towards an approaching vehicle, which has the right of way.

Scenarios which are resolved beforehand:

- The vehicle keeps its speed (or accelerates), the pedestrian slows down. Both driver and pedestrian non-verbally and mutually agreed that the vehicle passes first.
- The vehicle decelerates with the intention to stop before the pedestrian. This is perceived by the pedestrian, who keeps his pace (or accelerates) and crosses the road (sometimes thanking the driver and turning his head away from the vehicle)

Scenarios which typically lead to readjustments and – in some cases – explicit communication:

- The vehicle keeps its speed but the pedestrian does not slow down still looking at the vehicle. As this situation develops more critical the more time passes, at least one of the road users usually yields, letting the other one pass (resulting in the examples above).
- The vehicle decelerates, but so does the pedestrian. This potential "deadlock" situation usually results in the examples above (i.e. one of the TPs decelerating), with some sort of explicit communication by either road user. As the driver has the right of way but already decelerated, he usually will wave the pedestrian through if the velocity is low enough.

In the observed locations, either the pedestrian or the vehicle went first. This results in reoccurring patterns, which are depicted in the figures below:

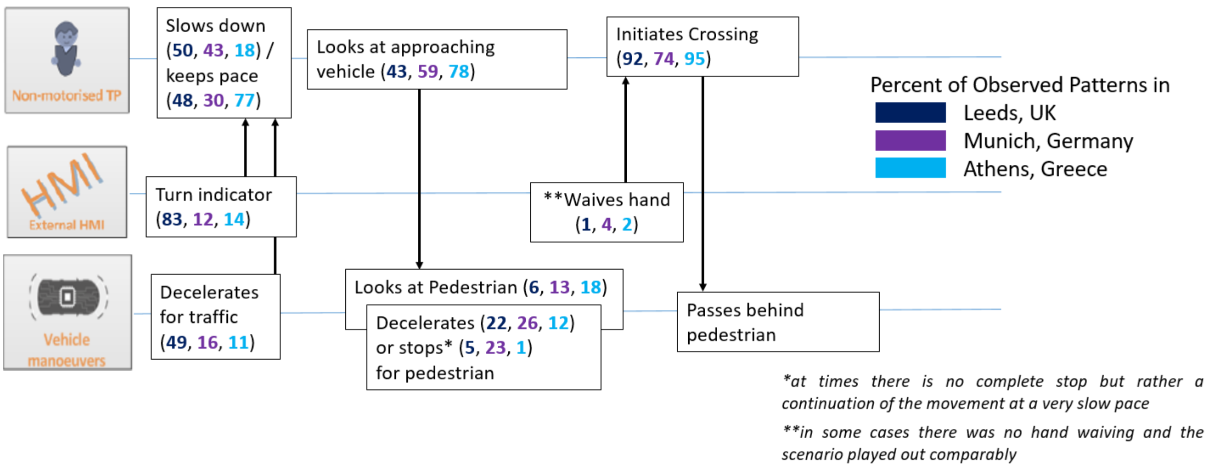


Figure 47: Sequence diagram of observed pedestrian vehicle encounters at intersections, where the pedestrian crosses in front of the vehicle. Numbers represent percentages of occurrences.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

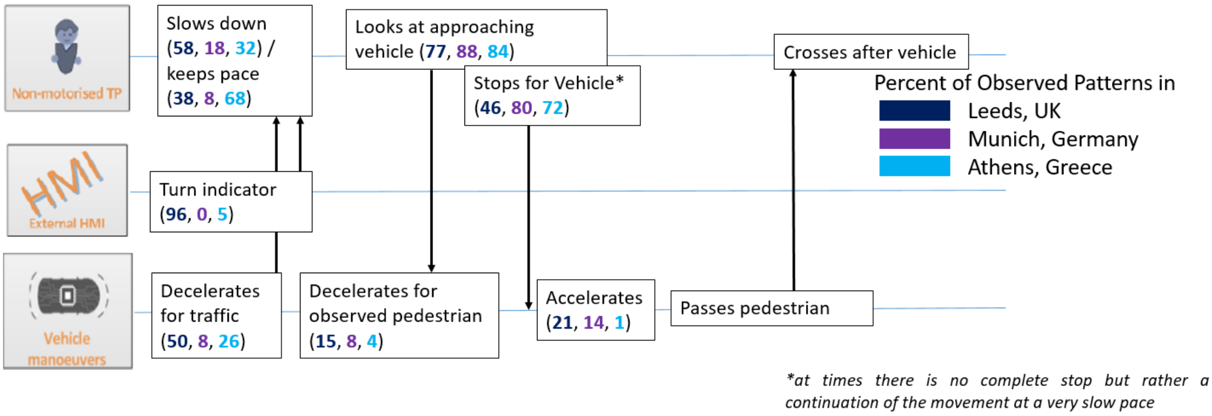


Figure 48: Sequence diagram of observed pedestrian vehicle encounters at intersections, where the pedestrian crosses after the vehicle passes. Numbers represent percentages of occurrences.

Within the shared space use cases, all traffic participants theoretically had the same priority. While this use case seems to encourage explicit communication, generally the situations played out comparably to the intersection: road users avoid communicating explicitly by adjusting their movements to resolve possible interaction-demanding situations early. E.g. if drivers see pedestrians walking on the right hand side they will adjust their lateral position towards the left and manoeuvre their vehicle around. Pedestrians usually indicate their intention to cross by turning and looking at an approaching vehicle – if said vehicle is close and keeps a lateral distance, pedestrians will cross after the vehicle. If the vehicle is further away, pedestrians will cross the road while drivers slow down or adjust their lateral position to the right.

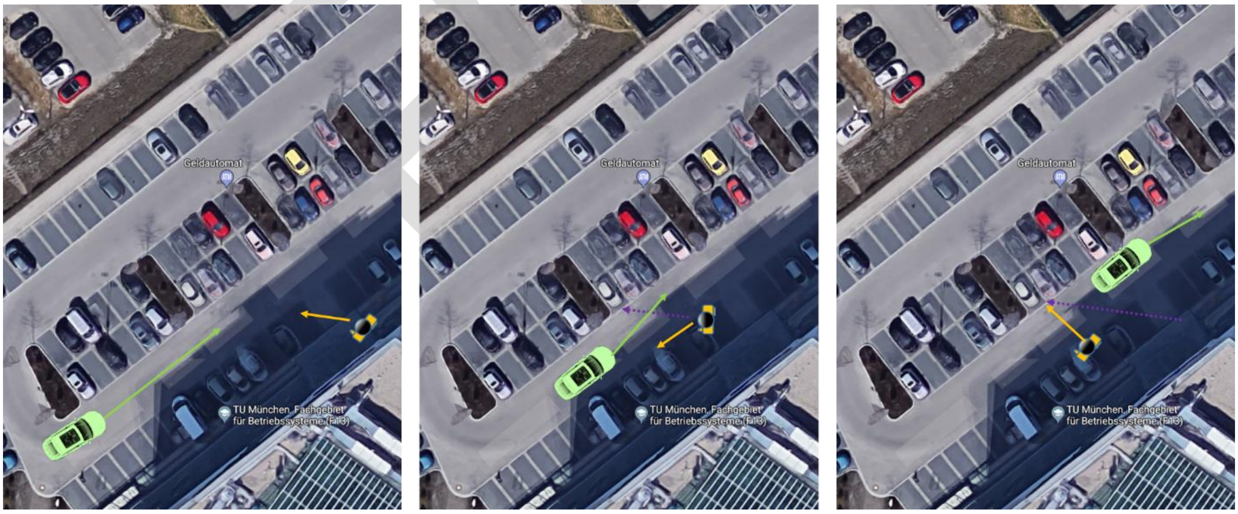


Figure 49: Depiction of a pedestrian-vehicle encounter on a shared space. The pedestrian aims to diagonally cross the road (dotted violet arrow), perceives the vehicle and decides to pass after it. The vehicle does not yield for the pedestrian, signalling this by keeping to its left.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

As a general recommendation for future AVs and the development of the CCPU based on the observations, we can say that a human-like, expectation conforming AV should try to avoid possible interaction-demanding situations or conflicts by adapting its driving behaviour early depending on the AVs intention in an underlying pedestrian-vehicle encounter (e.g. decelerate early to yield the right of way, or accelerate/keep the velocity to go first). Explicit communication using eHMIs should be used to resolve situations, in which the kinematic adaptation did not achieve the desired effect (e.g. deadlocks).

As AVs will be able to transmit signals before a driver could be fully perceived by a pedestrian, explicitly communicating a yielding behaviour early using eHMIs might see positive effects on pedestrians' crossing initiation and acceptance. Therefore, the general findings should not be treated as a universal guideline for an AV's behaviour but rather help in the development process of trajectory planning and communication capabilities. Furthermore, use cases different from those observed will potentially require other interaction strategies.

7.2 Vehicle-Vehicle encounters

In vehicle-vehicle encounters, observers perceived explicit communication more often. Usually these encounters turned into interaction-demanding situations, when one traffic participant tried to turn – either onto a priority lane or into a side road, crossing the oncoming lane. Traffic regulations stipulate that drivers on the priority road have the right of way while driving along the road. When turning, drivers have to indicate the upcoming manoeuvre and wait for a sufficiently large gap on the priority lane to turn. Road users usually follow these regulations, if the traffic density is low and uncongested, establishing rule based traffic conditions.

If prioritized road users would strictly insist on their right of way, the tailback could result in serious traffic jams. These situations usually have an increased traffic density, thus reducing the driven velocity on the priority lane. Some drivers on a congested priority lane yield their right of way for turning vehicles, as their progress and goal achieving is not considerably affected by letting another vehicle turn. There are different strategies, which drivers use to communicate their yielding behaviour – the reduction of the vehicle's velocity to create a gap was in some cases accompanied with either flashing the headlights, a waving hand gesture or nodding.

Furthermore, edging[8] into an intersection is an effective way to make the driver on the priority lane yield. This behaviour was observed in congested traffic situations, mostly in Greece, where turning drivers were waiting for gaps or yielding drivers, but to no avail. Edging was also observed in shared spaces, when drivers tried to pull out of parking spaces. Normally, other drivers would continue their movement increasing the lateral distance to the vehicle pulling out. Once the parking vehicle has backed out far enough (depending on the lane width), following drivers will yield and wait until the vehicle has left the parking space.

[8] Edging: moving forward with very low velocity usually to indicate a desired trajectory. Edging is mostly used by drivers, trying to pull out of a parking space with limited vision or while turning on congested priority lanes.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Typical sequences of interactions observed in Athens are depicted in Figure 50 and Figure 51.

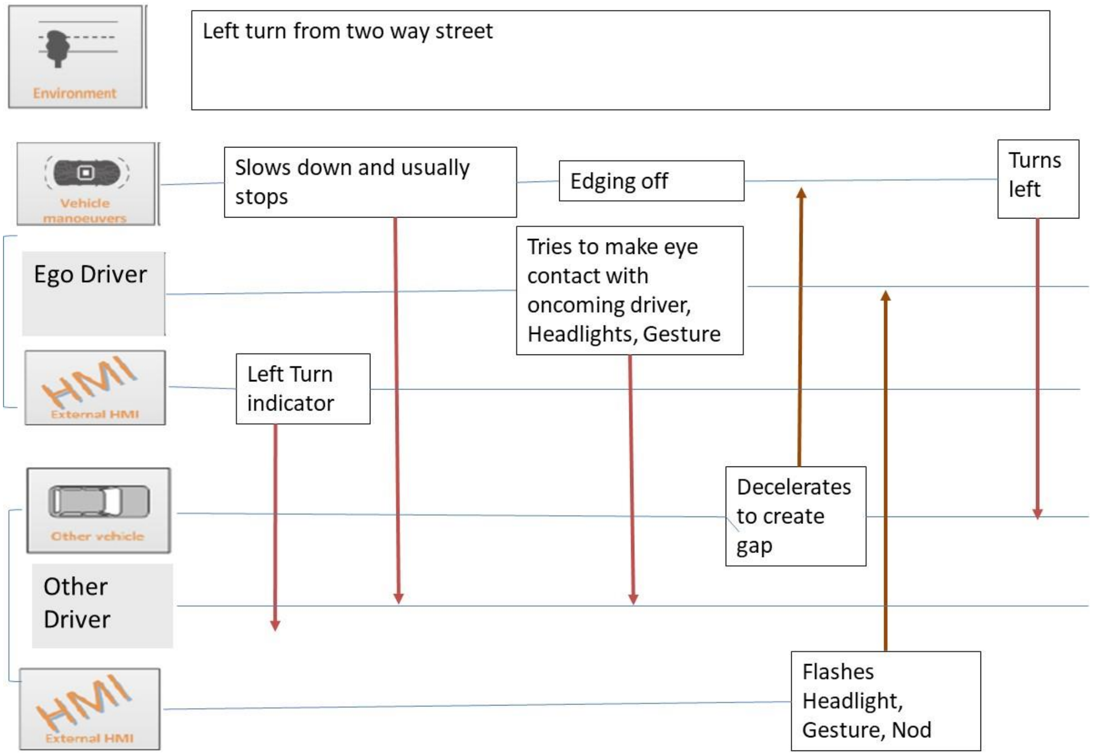


Figure 50: Typical interaction between drivers relevant to a left turn

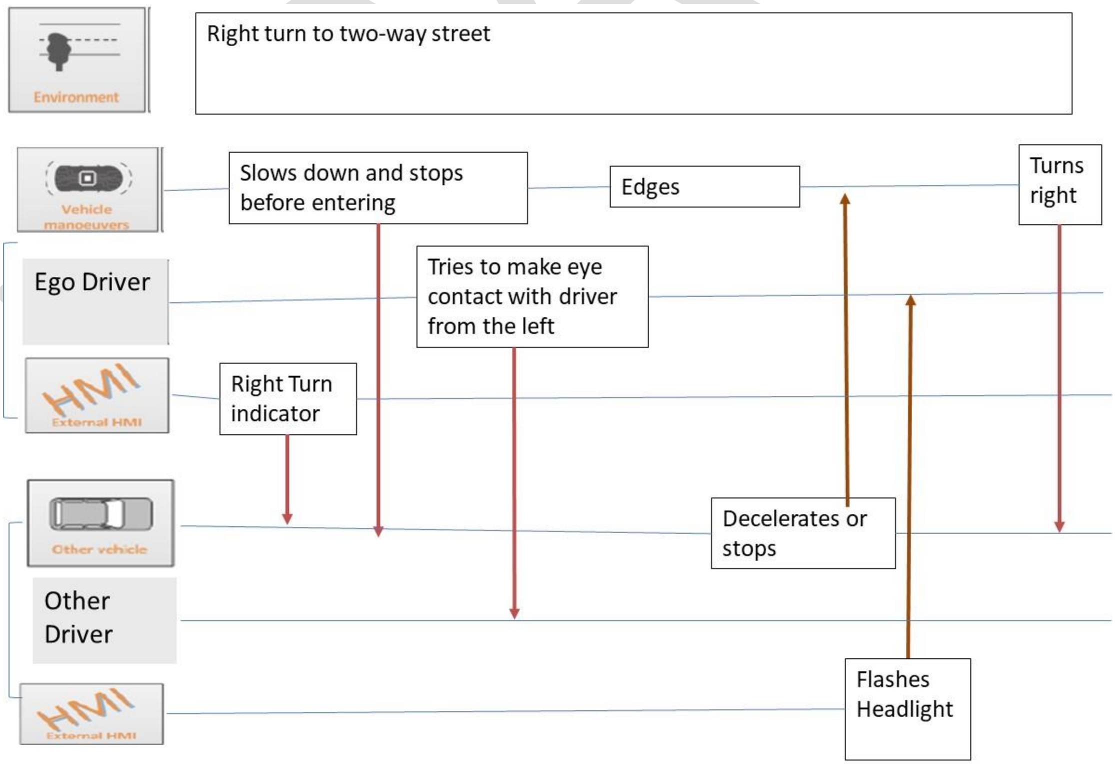


Figure 51: Typical driver-driver interaction before a right turn to two-way street

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Figure 52: Two different ways one particular use case can play out. Left: vehicle on the priority lane (blue) yields to let the other driver (purple) in, as traffic is congested in front. Right: driver on the side road edges into the intersection, forcing the driver on the priority lane to brake (red).

Example: A driver wants to turn onto a prioritized road, using the turn indicator and looking at the approaching traffic.

Normal traffic condition:

- Vehicles on the priority lane are keeping their speed and distances in between. The driver turning into the lane decelerates and brakes until standing still. When a sufficiently large gap occurs, the driver accelerates and turns onto the prioritized road.

Congested traffic condition, cooperative situation:

- Vehicles on the priority lane are driving way slower than usual due to the congestion. The traffic condition and low inter-vehicle distances do not allow the turning driver to merge. This is perceived by drivers on the priority lane, which will coast to increase the gap to the leading vehicle and – in some cases – explicitly signal the turning driver that they are letting them merge. The stopped driver turns onto the prioritized road usually thanking the yielding driver with a hand sign.

Congested traffic condition, enforced interaction:

- Vehicles on the priority lane are driving way slower than usual due to the congestion. The traffic condition and low inter-vehicle distances do not allow the turning driver to merge. As drivers on the priority lane are not letting the turning driver merge, he edges slowly into the intersection, still watching the approaching traffic. Once the driver edged his vehicle is almost in the path of the approaching traffic, usually some other vehicle will yield. In other cases the turning driver will merge into a small gap, forcing the other driver to decelerate. In both cases a “thank you” gesture is common.

This report is part of interACT project that has received funding from the European Union’s Horizon 2020 research and innovation programme under grant agreement no 723395.

AVs will encounter interaction-demanding situations from two perspectives in urban traffic – either on the road with right of way or when trying to turn onto one. Each situation requires different driving strategies, which an AV needs to adapt to.

Recommendations for turning onto a priority road:

If the AV tries to turn onto a priority road, it should wait for sufficient gaps or vehicles creating gaps by coasting/decelerating. If an approaching vehicle yields its right of way for the AV, it should merge edging into the intersection first and observe the other vehicle. If the approaching vehicle is not closing the gap by accelerating, the AV should merge rather quickly to not annoy the yielding driver, thanking him explicitly (e.g. using eHMIs). As urban traffic is very dynamic in congested situation, confidence is key – if the AV hesitates accepting a gap, the other driver might perceive this as a rejection and accelerate – potentially creating critical situations.

As AV's should prioritize safety over efficiency: it should not enforce interactions, as it would have to rely on the other driver to react appropriately. Therefore, in situations, where other drivers are not letting the AV turn, a take-over by the driver is probably necessary.

Recommendations for approaching a side road while on a priority road:

Human road users expect cooperation. Therefore, the AV should let a turning vehicle onto the priority lane, if the traffic flow is congested and multiple vehicles are behind the AV. To cooperate, the AV should coast early to increase the gap size in front and thereby inform the turning vehicle of the yielding behaviour. To avoid deadlocks, the AV should not fully stop but slowly roll into the intersection, giving the human driver enough time to merge. If the driver does not react until the (rolling) AV is about equidistant to the merging vehicle, it should accelerate to close the gap.

General recommendations for vehicle-vehicle encounters:

interACT should design explicit communication to other drivers, to attract their attention and to ensure that each other understands that they are each other's focal point. This could simulate the social conventions observed in the presented studies, where subjects felt that if they achieved eye contact, the other driver would feel obliged to react and yield.

An explicit signal by an automated vehicle to inform the other driver that the automated vehicle will yield/gives right of way may be beneficial for the traffic flow and efficiency.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

8. Summary & Outlook

This deliverable proposes a novel definition for interaction in urban traffic and gives an insight into a multicultural observation of urban traffic. Various methodologies, ranging from observing static locations using videos, protocols, questionnaires and a LiDAR to letting drivers comment on their behaviour after reviewing their own videos, were deployed to understand, how human road users behave in traffic conditions.

The procedures of the observations are described in detail, enabling future observational studies to incorporate the utilized methods. While each method yields a different data set, this deliverable creates a general understanding of current traffic for the use cases within D1.1.

The core finding of the observations is that urban traffic is not as interactive as one might think – most potential encounters are resolved by one traffic participant adapting to a present situation. Implicit communication, i.e. adjusting one's velocity to convey a certain behaviour, is prevalent in urban traffic interactions. Explicit communication is usually utilized additionally by drivers to either reassure another participant of a yielding intention, to resolve deadlocks or to warn another road user.

The observations are qualitatively modelled in sequence diagrams and will serve as an input for work packages 3 "Cooperation and Communication Planning Unit" and 4 "Suitable HMI for successful human-vehicle interaction".

The data analysis of the studies conducted within Work Package 2 will be pursued within Task 2.2 "Development of human-human and human-automation interaction models (qualitative and quantitative" to find answers for the open research questions presented in this deliverable. Furthermore, the effects of implicit yielding behaviour and explicit communication in general will be researched within T2.2 extending the sequence diagrams and models to create an understanding about future interactions with automated vehicles, which will be presented in deliverable D2.2 "Final description of psychological models on human-human and human-automation interaction".

Task 2.3 "Detecting interaction features and intention recognition development" aims to improve vehicle sensor algorithms to enable the perception and interpretation of the interaction vocabulary specified within this deliverable. Results of T2.3 will be described within deliverable D2.3 "Incorporation of sensors and algorithms for integration into the demonstrators".

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

9. References

Bogacz, R., Brown, E., Moehlis, J., Holmes P. and Cohen, J. D. (2006). The Physics of Optimal Decision Making: A Formal Analysis of Models of Performance in Two-Alterative Forced-Choice Tasks. *Psychological Review*. 113, 4, 700-765.

Chen, B., D. Zhao, and H. Peng. Evaluation of Automated Vehicles Encountering Pedestrians at Unsignalized Crossings. https://arxiv.org/abs/1702.00785.

Clamann, M., Aubert, M., & Cummings, M.L. (2017). Evaluation of vehicle-to-pedestrian communication displays for autonomous vehicles. *Paper presented at the Transportation Research Board 96th Annual Meeting,* Washington DC, United States, 8th-12th January 2017.

Dalal, N. and Triggs, B., "Histograms of oriented gradients for human detection," *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA, USA, 2005, pp. 886-893 vol. 1.

Díaz, E. M. (2002). Theory of planned behavior and pedestrians' intentions to violate traffic regulations. *Transportation Research Part F*, 5, 169-175.

Elliott, M.A. & Baughan, C.J. (2004). Developing a self-report method for investigating adolescent road user behaviour. *Transportation Research Part F,* 7, 373-393.

Ericsson, K.A. and Simon, H.A. (1993). Protocol analysis: Verbal reports as data, *MIT Press, Cambridge, MA*

Fuest, T., Sorokin, L., Bellem, H., & Bengler, K. (2017). Taxonomy of traffic situations for the interaction between automated vehicles and human road users. *International Conference on Applied Human Factors & Ergonomics,* 708-719, Cham: Springer.

Guéguen, N., Meineri, S., & Eyssartier, C. (2015). A pedestrian's stare and drivers' stopping behavior: a field experiment at the pedestrian crossing. Safety Science, 75, 87–89.

Hamed, M.M. (2001). Analysis of pedestrians' behaviour at pedestrian crossings. *Safety Science,* 38, 63-82.

Hanheide, M. & Dondrup, C., (2014, June 5). First standalone release. Zenodo. http://doi.org/10.5281/zenodo.10318

Harrell, W. A. (1991). Factors influencing pedestrian cautiousness in crossing streets. *Journal of Social Psychology*, 131, 367–372.

Ishaque, M.M., & Noland, R.B. (2008). Behavioural issues in pedestrian speed choice and street crossing behaviour: A review. *Transport Reviews,* 28 (1), 61-85.

Kesting, A., M. Treiber, M. Scho, et al. Adaptive cruise control design for active congestion avoidance. *Transportation Research Part C*, 16, 668–683. http://dx.doi.org/10.1016/j.trc.2007.12.004.

Klop, J.R., & Khattak, A.J. (1999). Factors influencing bicycle crash severity on two-lane, undivided roadways in North Carolina. *Transportation Research Review*, 1674, 78–85.

Luo, W.; Zhao, X. & Kim, T. Multiple Object Tracking: A Review, CoRR, 2014, abs/1409.7618

Marisamynathan, S., & Vedagiri, P. (2013). Modeling pedestrian delay at signalized intersection crosswalks under mixed traffic condition. *Procedia – Social and Behavioural Sciences,* 104, 708-717.

Markkula, G., J. Engström, J. Lodin, et al (2016). A farewell to brake reaction times? Kinematics-dependent brake response in naturalistic rear-end emergencies. *Accident Analysis & Prevention*, 95, 209–226.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Markkula, G., Romano, R., Madigan, R., Fox, C. W., Giles, O. T., & Merat, N. (2018). Models of human decision-making as tools for estimating and optimising impacts of vehicle automation. Transportation Research Record, in press, http://journals.sagepub.com/home/trr.

Miletic, S., Turner, B. M., Forstmann, B. U. and van Maanen, L. (2017). Parameter recovery for the Leaky Competing Accumulator model. *Journal of Mathematical Psychology*. 76, 25-50.

Oxley, J.A., Ihsen, E., Fildes, B.N., Charlton, J.L., & Day, R.H. (2005). Crossing roads safely: An experimental study of age differences in gap selection by pedestrians. *Accident Analysis & Prevention*, 5, 962-971.

Özlu, A. Vehicle Detection, Tracking and Counting by TensorFlow, 2018, https://github.com/ahmetozlu/vehicle_counting_tensorflow

Özlü, A., Vehicle Detection, Tracking and Counting, 2017, https://github.com/ahmetozlu/vehicle_counting_hog_svm

Portouli, E., Nathanael, D., Marmaras, N. (2014). Drivers' Communicative Interactions: On-road observations and modelling for integration in future automation systems. *Ergonomics,* DOI: 10.1080/00140139.2014.952349

Ratcliff, R., Smith, P., Brown, S. and McKoon, G. (2016). Diffusion Decision Model: Current Issues and History. *Trends in Cognitive Sciences*, 20, 4, 260-281.

Rosenbloom, T., Nemrodov, D., & Barkan, H. (2004). For heaven's sake follow the rules: pedestrians' behavior in an ultra-orthodox and a non-orthodox city. *Transportation Research Part F: Traffic Psychology and Behaviour,* 7, 395–404.

Rudra Gouda, J. S. N. Survey on Pedestrian Detection, Classification and Tracking *(IJCSIT) International Journal of Computer Science and Information Technologies*, 2015, 6 (4)

SAE International Surface Vehicle Recommended Practice, "Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles", SAE Standard J3016, Rev. Sep. 2016.

Sayed, T., Zaki, M.H., & Autey, J. (2013). Automated safety diagnosis of vehicle–bicycle interactions using computer vision analysis. *Safety Science*, 59, 163–172

Sobral, A., BSLibrary: An OpenCV C++ Background Subtraction Library, *IX Workshop de Visão Computacional (WVC'2013)*, Rio de Janeiro, Brazil, 2013, June, https://github.com/andrewssobral/bgslibrary

Sobral, A., Vehicle Detection, Tracking and Counting, 2014, https://github.com/andrewssobral/simple_vehicle_counting

Šucha, M. (2014). Road users' strategies and communication: Driver pedestrian interaction. Proceedings 5th Conference Transport Solutions from Research to Deployment, Transport Research Arena, Paris.

Svärd, M., G. Markkula, J. Engström, et al (2017). A quantitative driver model of pre-crash brake onset and control. *Presented at Human Factors and Ergonomics Society 61st Annual Meeting*.

Toni, T., Welch D., Strekowa N., Ipsen, A., Stumpf, M. P. H (2009). Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of the Royal Society Interface*. 6, 31, 187-202.

Wagner, J. (1981). Crossing streets: Reflections on urban pedestrian behavior. *Man-Environment Systems*, 11, 57-61.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Zhou, R., Horrey, W.J., & Yu, R. (2009). The effect of conformity tendency on pedestrians' road-crossing intentions in China: an application of the theory of planned behavior. *Accident Analysis & Prevention,* 41, 491–497.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Annex 1: Questionnaires

1.Participant No. ____________________

2. Age ______________________________________

3. Gender ___________________________________

4. Travelling From:

(Mark with X in 1 of the boxes below)

Home	**Shopping**	**Work/Education**	**Visiting Friend**	**Other**

5. Going to:

(Mark with X in 1 of the boxes below)

Home	**Shopping**	**Work/Education**	**Visiting Friend**	**Other**

6. How regularly do you use this crossing?

(Mark with X in 1 of the boxes below)

Daily	
More than once a week	
Weekly	
More than once a month	
Monthly	
Less often	
First time	

7. How safe did you feel during that crossing?

(Mark with X in 1 of the boxes below)

Very unsafe	**Unsafe**	**Safe**	**Very safe**

Why?

____

____

8. What information from the vehicle, if any, did you use to decide it was safe to cross?
(Mark with X in 1 or more of the boxes below)

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Speed	
Distance	
Braking	
Flashing Lights	
Vehicle Trajectory	
Turn Indicator	
Waited for vehicle to pass	
None	

9. What information from the driver did you use to decide it was safe to cross? (Mark with X in 1 or more of the boxes below)

Watching driver	
Mutual Eye contact	
Hand gesture	
Head nod	
Head movement to the side	
None	

Was there any other information you used to determine how safe it was to cross?

(To be completed if None is selected in response to Q8 or Q9)

____

____

10. How long did you feel you were waiting to find a suitable crossing gap?

Longer than usual	
Average	
Shorter than usual	

11. Did the presence of other people affect your decision of when to cross?

Yes ☐ No ☐

- If so, in what way?

____

____

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

12. How did you indicate your intention to cross the road?

Stepping forward	
Eye contact	
Hand gesture	
Head movement (Looking around)	
Other	
None of the above	

13. Who do you think had priority in this situation?

You ☐ The driver ☐

14. Are you a car driver?

Yes ☐ No ☐

15.

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Adolescent Road User Behaviour Questionnaire (Elliott & Baughan, 2004)

How often do you…	Never	Rarely	Some-times	Often	Very Often
Forget to look properly because you are talking to friends who are with you					
Cross from between parked cars when there is a safer place to cross nearby					
Think it is OK to cross safely, but a car is coming faster than you thought					
Forget to look properly because you are thinking about something else					
See a small gap in traffic and ‘‘go for it’’					
Run across a road without looking because you are in a hurry					
Cross whether traffic is coming or not, thinking the traffic should stop for you					
Get part way across the road and then have to run the rest of the way to avoid traffic					
Cross from behind a stationary vehicle					
Cross when you cannot see both ways very well (like on a bend or top of hill)					
Not look because you cannot hear any traffic around					
Use a mobile phone and forget to look properly					
Not notice a car pulling out (say from a driveway) and walk in front of it					
Cross without waiting for the ‘‘green man’’					
Climb over barriers or railings that separate the road from the pavement					
Walk in the road rather than on the pavement					

This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Annex 2: Observation Protocol and Screenshots from the Observation App

Excel based protocol:

Vehicle - Vehicle - Interaction Observation Protocol

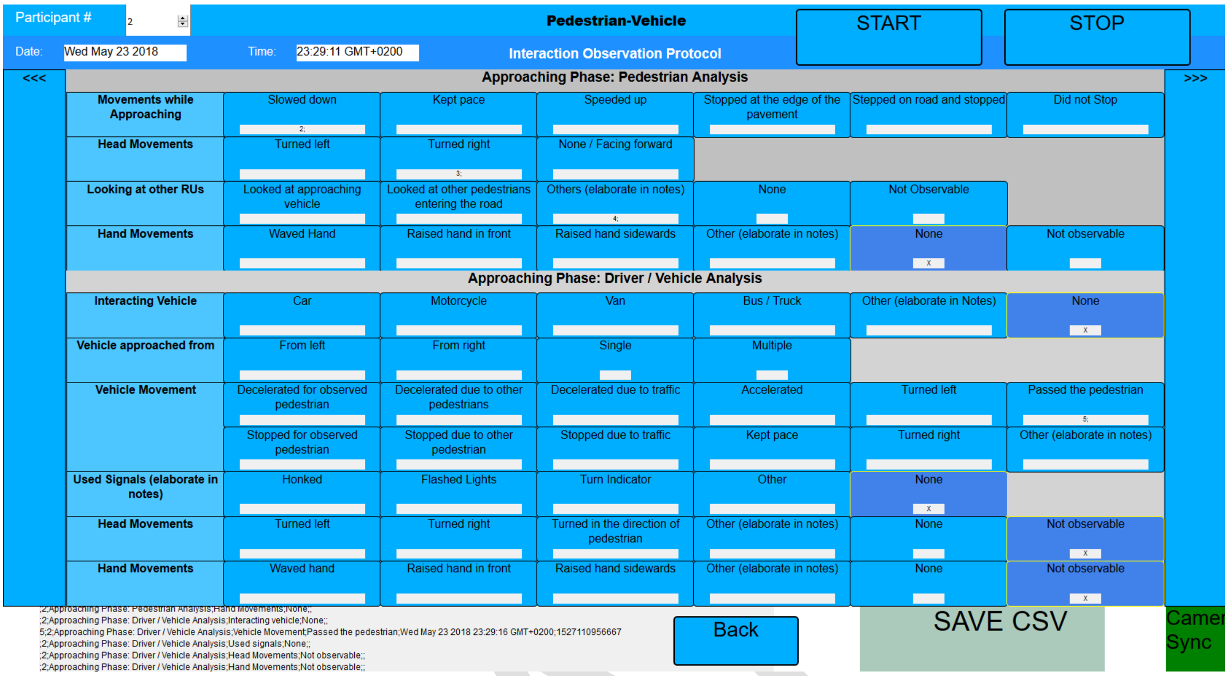
Vehicle 1 Analysis = The vehicle that intends to complete a certain action (e.g. turning, crossing, etc.)						
Vehicle Movement	Decelerated for vehicle 2	Decelerated due to traffic	Kept pace	Crept into the intersection	Turned left	Passed vehicle 2
	Stopped for vehicle 2	Stopped due to traffic	Accelerated	Entered intersection first	Turned right	Other (elaborate in notes)
Used signals add the meaning of used signals in notes section	Honked	Flashed Lights	Turn Indicator	Other (elaborate in notes)	None ☐	
Head movements add meaning of hand movement in notes section	Turned left	Turned right	Turned in direction of vehicle 2	Turned in the direction of pedestrians	None/ Facing forward ☐	Not observable ☐
Hand movements add meaning of hand movement in notes section	Waved hand	Raised hand in front	Raised hand sideways	Other (elaborate in notes)	None ☐	Not observable ☐
Vehicle 2 Analysis						
Vehicle Movement	"Decelerated for vehicle 1	Decelerated due to traffic	Kept pace	Crept into the intersection	Turned left	Passed vehicle 1
	Stopped for vehicle 1	Stopped due to traffic (elaborate in notes)	Accelerated	Entered intersection first	Turned right	Other (elaborate in notes)
Used signals add the meaning of used signals in notes section	Honked	Flashed Lights	Turn Indicator	Other	None ☐	
Head movements add meaning of hand movement in notes section	Turned left	Turned right	Turned in direction of vehicle 1	Turned in the direction of pedestrians	None/ Facing forward ☐	Not observable ☐
Hand movements add meaning of hand movement in notes section	Waved hand	Raised hand in front	Raised hand sideways	Other (elaborate in notes)	None ☐	Not observable ☐
Date: **Intersection:**						
Weather	Sunny ☐	Overcast ☐	Raining ☐	Freezing / Icy ☐		
Vehicle 1	Car ☐	Motorcycle ☐	Van ☐	Bus / Truck ☐	Cyclist ☐	Other (elaborate in notes) ☐
Vehicle 2	Car ☐	Motorcycle ☐	Van ☐	Bus / Truck ☐	Cyclist ☐	Other (elaborate in notes) ☐
Vehicle 2 approached vehicle 1 from	From left	From right	From front	From behind		

Additional Observation Notes:

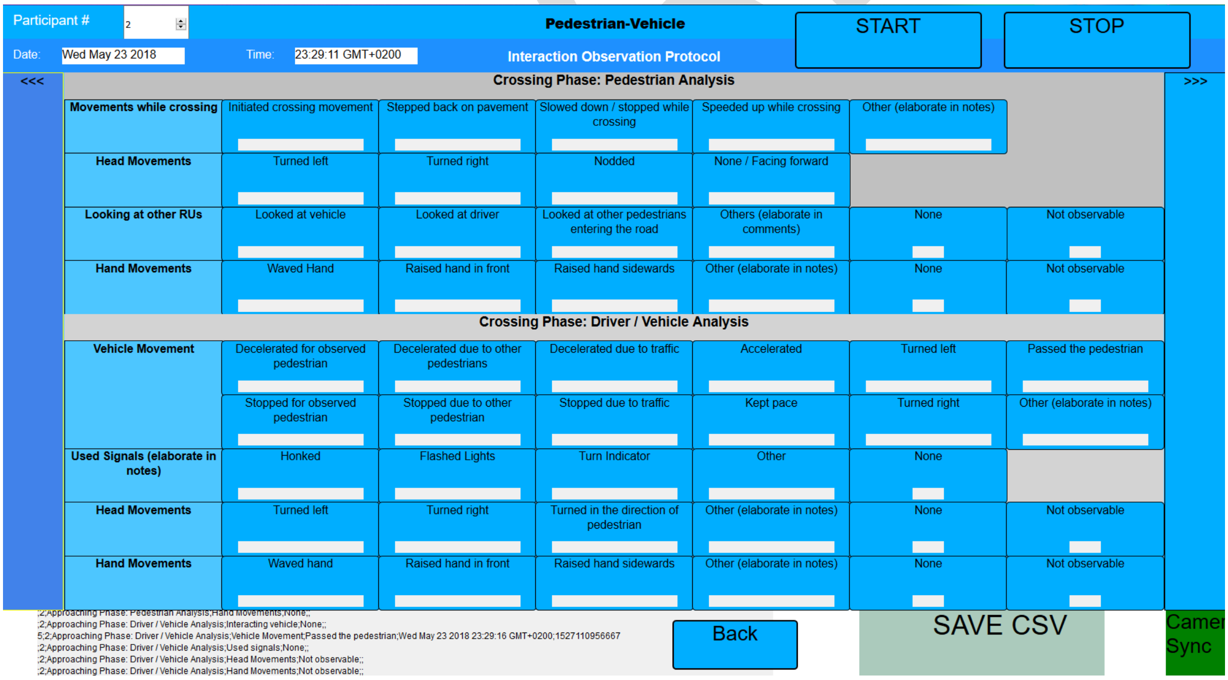
This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Interact App: Protocol for pedestrian-vehicle interactions

Page 1 – Approaching Phase

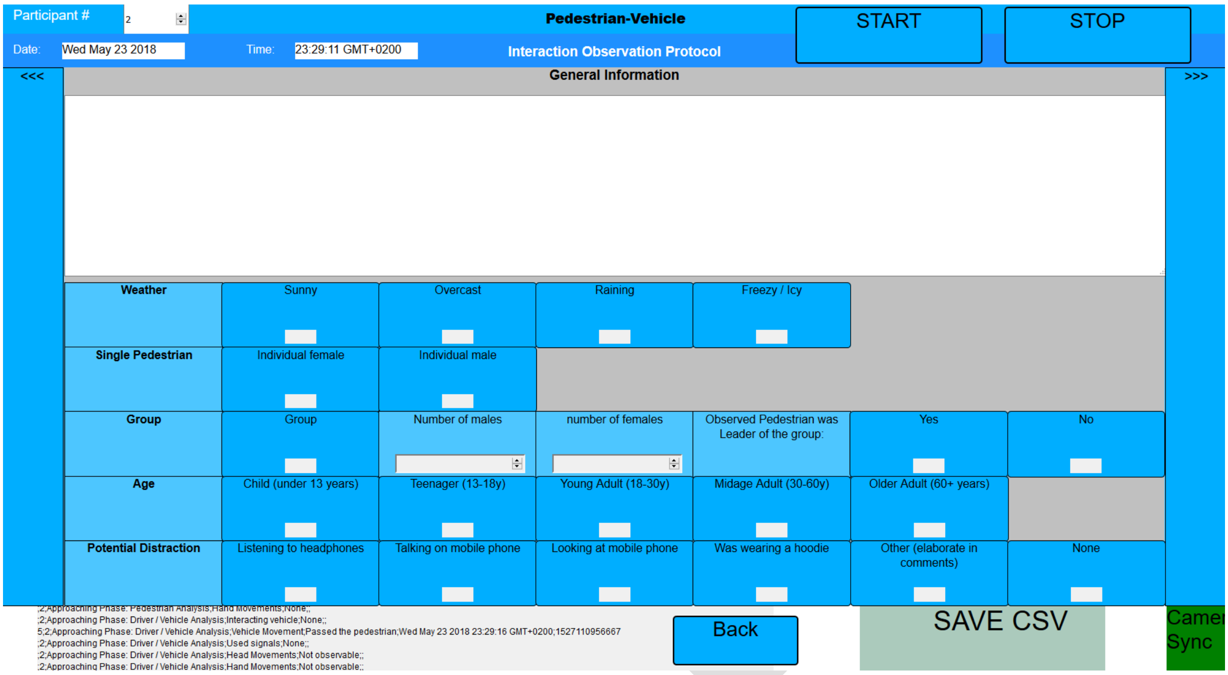


Page 2 – Crossing Phase

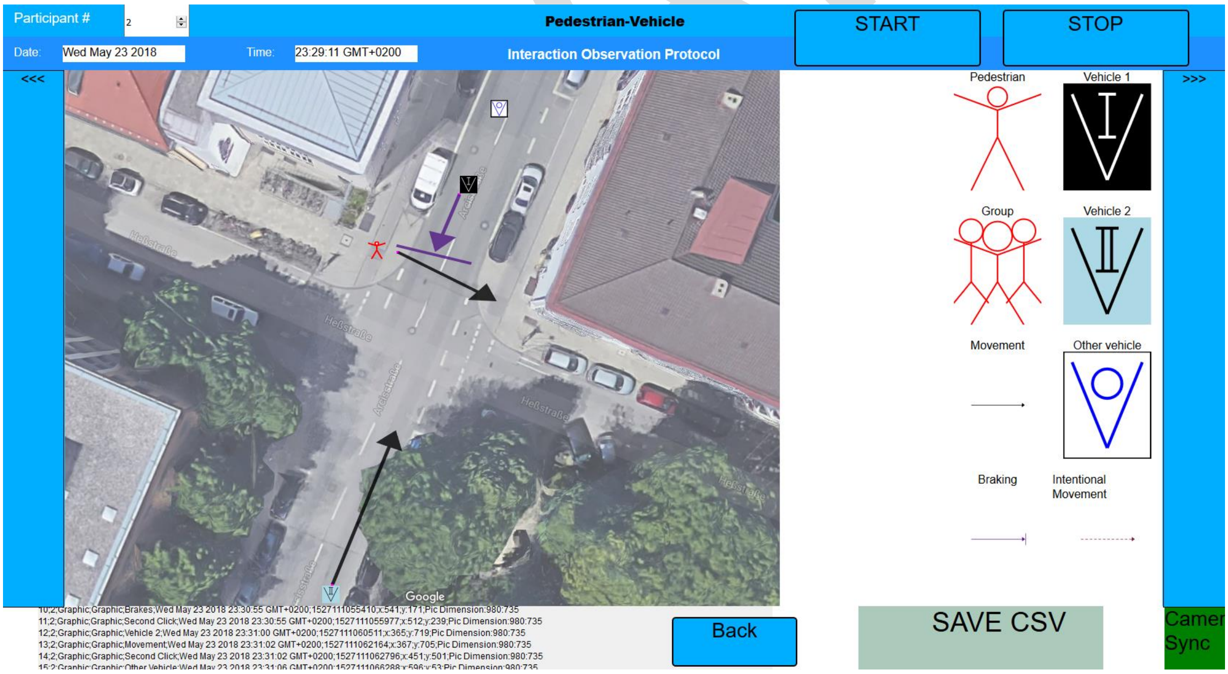


This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

Page 3 – General Information



Page 4 – Depiction of the pedestrian-vehicle interaction



This report is part of interACT project that has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no 723395.

For more information:

interACT Project Coordinator

Anna Schieben

DEUTSCHES ZENTRUM FUER LUFT - UND RAUMFAHRT e.V. (DLR)

Lilienthalplatz 7

38108 Braunschweig, Germany

Anna.Schieben@dlr.de

interact-roadautomation.eu/

Designing cooperative interaction of automated vehicles with other road users in mixed traffic environments